IT University of Copenhagen

Computer Science Department

Ph.D. Programme in Computer Science

Using Priors to Improve Head-Mounted Eye Trackers in Sports

Ph.D. Thesis

By:
Fabricio Batista Narcizo

Supervisor:
Dan Witzner Hansen

Copenhagen – Denmark

November, 2017

Using Priors to Improve Head-Mounted Eye Trackers in Sports

by

Fabricio Batista Narcizo

B.Sc. University of the West of Santa Catarina, 2005
M.Sc. Technological Institute of Aeronautics, 2008

A thesis submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Computer Science Department
IT University of Copenhagen

November, 2017

Signature of the Author _____

Accepted by _____
Coordinator, Ph.D. Programme Date

IT University of Copenhagen
Computer Science Department
Ph.D. Programme in Computer Science
Copenhagen – Denmark

Certificate of Approval

Ph.D. Degree Dissertation

The Ph.D. Degree Dissertation of Fabricio Batista Narcizo has been examined and approved by the dissertation committee as satisfactory for the dissertation required for the Ph.D. degree in Computer Science

---

Dr. Dan Witzner Hansen, Dissertation Supervisor

Dr. Björn Thór Jónsson, Chair

Dr. Arantxa Villanueva

Dr. Ben William Tatler

Date

This Ph.D. thesis is dedicated to my entire family with love, specially to my father José Narcizo Sobrinho. Actually Dad doesn't read anything related to Computer Science, so if someone doesn't tell him about this, he'll never know.

Acknowledgments

Gratitude is one of the noblest feelings of a man. Thus, I could not fail to mention that the accomplishment of this work is the result of the collaboration of many people. I would like to express my gratitude in a particular way.

First to God, my Father, Director, and Author of my life, Who gave me strength when I missed my relatives and friends. Without Him, I would not be writing these words of thanks. Because, for Him are all things, including my life.

I thank my supervisor Dan Witzner Hansen, who gave me the opportunity to study and work in this institution, providing valuable guidance on academic paths. Furthermore, Dan is a friend and a confidant, and through his confidence in me, I have gained wisdom.

I thank the IT University of Copenhagen (ITU) for the opportunity to have participated in this family, including all teachers and staffs. The Brazilian National Council for Scientific and Technological Development (CNPq) for the academic financial support to my research project. The Bandeirante University of São Paulo, the University July 9^{th}, the Mackenzie Presbyterian University, and the Federal University of Campina Grande that have believed in my academic potential.

I also thank my former supervisors Lilian Jeannette Meyer Riveros (UNOESC), Carlos Henrique Quartucci Forster (ITA), Felipe Afonso de Almeida (*in memoriam*) (ITA), Herman Martins Gomes (UFCG), and José Eustáquio Rangel de Queiroz (UFCG), for making me a better man, researcher, student, and person.

Special thanks to the masters Carlos Henrique Costa Ribeiro (ITA), Nei Yoshihiro Soma (ITA), Sérgio Roberto Matiello Pellegrino (ITA), Leila Maria Garcia Fonseca (INPE), João Ricardo de Freitas Oliveira (INPE), Jaime de Britto (UNOESC), Eugênio Simão (UNOESC) and Ernani Tadeu Rizzi (UNOESC), who passed on valuable knowledge

during my academic life. They are my sources of inspiration and my greatest teachers.

I could not forget some students that become good friends. Edivaldo Santos da Cunha, Marcos Velloso, Danielle Katucha Garcia Moreno, Celso do Amaral, Gabriel Henrique Alves Marin, Leonard Colusso Oliveira, Verônica Longuinho Silva de Magalhães, Tainã Miranda de Souza, Maytham Mohamed Fahmi, Martin Hahner, and Ada Maria Lorenc.

Blessed are those who have friends, and I take with me the friendship of incredible people from ITU (in alphabetic order): Amanuel Zewge, Andres Faina, Carla Ysabela Villegas Pasco, Ceyue Liu, Diako Mardanbegi, Farzad Nejatimoharrami, Florian Biermann, Isabel Cristina Gonçalves Fróes, Ivan Luiz Picoli, Jean Carlos de Carvalho Melo, Jonathan Fürst, Joshua Ddamba, Luiz Adolfo Andrade, and Paolo Tell.

The greatest thanks go to Zaheer Ahmed, my friend, my colleague and for last two years my research partner. I have to thank his professional and energetic support helping me in the development and evaluation of the EyeInfo Framework, tasks which I could not have achieved alone. A particular thanks to Otto-Max Klein, Melanie Mack and Jean Carlos de Carvalho Melo for the collaboration between our projects. And a "*mange tak*" to Isabel Cristina Gonçalves Fróes for helping me in my abstract in Danish.

I also thank my pet Chuck Norris. He gave me an extra strength so I could finish this research. Sometimes, taking care my family, drawing my attention, relieving the stress of work, and other times, listening intently to rehearsals for my Ph.D. defense. Thanks Fióte. -Give a Roundhouse Kick on Egídio!

I would like to thank the Munzlinger family for supporting me along the last 16 years, in special to Mr. Pedro Munzlinger and Mrs. Anaides Sauthier Munzlinger.

To my parents José Narcizo Sobrinho and Rita de Cássia Soares Batista Narcizo, my sister Rita de Cássia Batista Narcizo, my bother-in-law Raydem Rabello Santana, and my niece Rafaela Narcizo Rabello for their continuous and unparalleled love, help and support.

My wife, Elizabete Sauthier Munzlinger Narcizo has been extremely supportive of me throughout this entire process and has made countless sacrifices to help me get to this point. I love you!

Finally, my last thanks goes to the most important person in my life. My angel, my little princess, my daughter Maria Rita Munzlinger Narcizo. I simply love you, more than I love life itself!

"With regard to performance, commitment, effort,
dedication, there is no middle ground.
Or you do something very well or not at all."

Ayrton Senna do Brasil (1960-1994)

Abstract

This Ph.D. thesis is about using available information known from the problem at hand (aka *priors*), with the aim to enhance the performance of head-mounted eye trackers. Prior information is used for eye tracking scenarios in different sports disciplines to improve the accuracy and robustness of gaze estimation in critical situations. This thesis also explores off-the-shelf hardware to build flexible and adaptable eye trackers that exploit the constraints revealed for specific sports settings. Several eye tracking methods are presented, in which the use of priors plays the leading role. The compensation models proposed in this thesis ranging from solving geometrical constraints of head-mounted eye trackers to eye feature detection in challenging environment lighting conditions. The experiments focused on different sports disciplines to collect and analyze eye tracking data involving elite athletes during the daily training sessions of shooting and kayak as well as some laboratory experiments. The results of the experiments showed that the use of priors is very promising to the field of eye tracking, such as (i) using the distance between the athlete and the observed target as priors, to reduce the influence of parallax error in 80.59%; (ii) using the 3D angles from the athlete's head as priors, to reduce the influence of head rotation in 86.41%; (iii) using the geometric relation of human ocular system as priors, to make eye tracking more robust to eye feature noise, among others. Using priors in different steps of an eye tracking system has a general and substantial impact on eye trackers in general. While the focus of this thesis is in the use of eye tracking in sports, it is evident that progress achieved within this project on gaze estimation for sports activities has a direct impact on other areas that use eye tracking as well.

Keywords: eye tracking, sports analysis, prior information, head-mounted eye tracker.

Resume

Denne Ph.D. afhandling handler om at anvende tilgængelig information vedrørende det aktuelle problem (aka *priors*), med det formål at forbedre ydeevnen i hovedmonterede eye trackers. Priors anvendes til eye tracking-scenarier inden for forskellige sportsdiscipliner med henblik på at forbedre nøjagtighed og robusthed i blikbedømmelsen i kritiske situationer. Denne afhandling udforsker også tilgængelig hardware til opbygningen af fleksible og tilpasningsdygtige eye trackers, der udnytter de begrænsninger, som har vist sig inden for bestemte sportsindstillinger. Flere forskellige metoder til eye tracking præsenteres, hvor brugen af priors spiller en afgørende rolle. De kompensationsmodeller, der foreslås i denne afhandling, spænder fra at løse geometriske begrænsninger i hovedmonterede eye trackers til detektering af øjetræk under vanskelige belysningsforhold. Eksperimenterne fokuserede på forskellige idrætsdiscipliner i indsamlingen og analysen af eye tracking-data under eliteatleters daglige træningspas inden for skydning og kajak samt laboratorieforsøg. Resultaterne af eksperimenterne viste, at brugen af priors er meget lovende for eye tracking-området, såsom (i) anvendelse af afstanden mellem atleten og det observerede mål som prior for at reducere indflydelsen af paralaksefejl i 80,59% ; (ii) at bruge 3D-vinklerne fra atletens hoved som prior for at reducere indflydelsen af hovedrotation i 86.41%; (iii) Brug af det geometriske forhold i det menneskelige okulære system som priors for at gøre eye tracking mere robust i forhold til blandt andet øjetræk-støj. Brug af priors på forskellige trin i et eye tracking-system har en generel og betydelig indvirkning på eye tracking generelt. Selv om fokus i denne afhandling er anvendelsen af eye tracking til analyse inden for sport, er det tydeligt, at de fremskridt der er opnået inden for dette projekt vedrørende sportsaktiviteter, kan have en direkte indvirkning på andre områder, hvor eye tracking også anvendes.

Nøgleord: eye tracking, sportsanalyse, prior information, hovedmonteret eye tracker.

Contents

Part I

Introduction and Foundation

Chapter 1

Introduction

This Ph.D. thesis presents the use of available information known from the problem at hand (aka *priors*), with the aim to enhance the performance of head-mounted eye trackers to be used in sports. *Prior information* is used to solve classical problems highlighted in the literature, such as (i) using the known distance between the subject and the observed target as priors, to compensate the parallax error [1] in head-mounted eye trackers (see Chapter 4); using 3D angles from the user's head as priors, to compensate the influence of head rotations [2] in gaze estimation (see Chapter 5); using the geometric relationships between both eyes as priors, to validate a binocular eye feature detector in outdoor conditions [3] (see Chapter 6), among others.

The project was inspired in different sports disciplines to assist and evaluate the development of accurate and adaptable eye tracking systems. The use of head-mounted eye trackers in the wild is still a challenging research topic in the field of eye tracking, especially due to problems in environment light conditions, geometric constraints in uncalibrated eye trackers, and complicated setup of the eye tracking components. This Ph.D. thesis aims (i) to investigate methods and mathematical models to enhance robustness in eye tracking; (ii) to explore off-the-shelf hardware to build flexible and fitting eye tracking equipment; (iii) to develop new methods for eye tracking that exploits the constraints revealed for specific sports settings; and (iv) to use prior information to improve the robustness and accuracy of eye tracking systems.

1.1 Overview

Eye tracking has been showing a great potential for detailed and objective performance analysis in sports experiments. Most scientific research that uses eye tracking in sports

is of psychological nature [4, 5, 6, 7], but eye tracking also provides detailed performance measures of elite athletes that cannot be obtained otherwise [8, 9, 10, 11].

Elite athletes need to perform optimally. In many sports disciplines, it is a matter of being fast, precise and accurate at the same time. When the ocular activities are the primary skill of some sport (e.g. shooting [5]), the athlete must look at the right thing at the right time to achieve the best performance. Detailed eye tracking data allow the athletes and trainers to get much deeper insight into strategies used by the athletes during the fulfillment of a sport action [8, 9, 12]. They can also adapt the training correspondingly, thus improving their performance in stressful and time critical situations [10]. Furthermore, novice and intermediate athletes can use eye information for quickly acquiring knowledge that expert athletes may not even be explicitly known [9, 13].

Gaze data collected during daily training sessions or competitions are usually used to answer empirical questions, such as (i) what do athletes look at during specific sports actions? (e.g. kick, jump, leap, punch, shot, catch, throw, attacking, defending) [8, 9, 12] (ii) is there any difference in the ocular activities of novice, intermediate and experts athletes? [8, 10, 12, 14] (iii) which strategies can be used to improve the novice athletes' performance based on knowledge of eye movements patterns of expert athletes? [9, 13] (iv) which visual information are athletes gathering? [4, 5] (v) which are the areas of interest viewed during sports practices? [11] (vi) how long do athletes focus on areas of interest? [11] and (vii) when does gaze behavior (e.g. fixation, saccades, blinks, dilation, quite eye, onset, offset) occur? [7, 10, 14, 15, 16].

In general, athletes perform sports activities in highly dynamic environments that are difficult to control. Thereby, the monitoring of eye information from elite sports athletes needs to be accomplished by specialized tools and devices.

1.2 Problem Delimitation

Hansen and Ji [17] describe fundamental eye and gaze tracking models that can be used to build eye trackers for sports experiments. Even nowadays, accuracy, precision and robustness of head-mounted eye trackers are still negatively influenced by problems such as (i) parallax error [1, 18, 19, 20, 21, 22], (ii) head rotations [2, 23, 24, 25], and (iii) camera slippage [26, 27, 28, 29]. In general, several kinds of noises from the environment influence the detection of eye features in video-based eye tracking. For example, eye tracking in outdoor conditions (a common practice in many sports disciplines) is difficult and

challenging due to sunlight, eye occlusions, reflections, small pupil circumference, low contrast between pupil and iris, among others [3, 30, 31, 32, 33, 34], as shown in Figure 1.

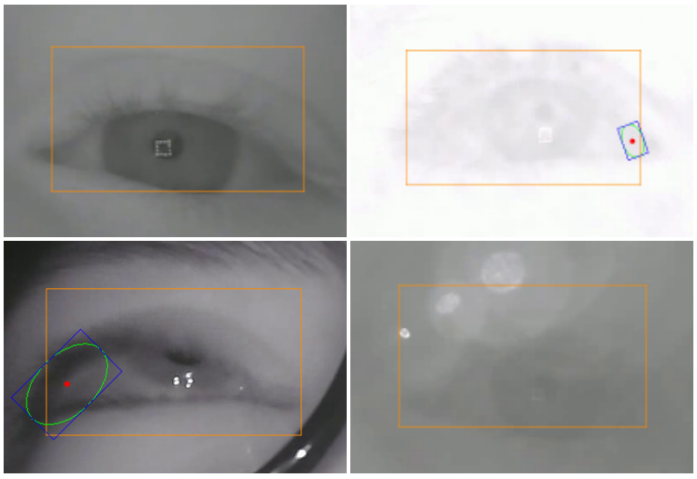


Figure 1: Pupil detection usually fail due to (*top-left*) reflections over the pupil; (*top-right*) illumination condition changes; (*bottom-left*) eye occlusions; and (*bottom-right*) low contrast between iris and pupil.

Most current scientific studies of eye tracking in sports have used head-mounted eye trackers [8, 9, 10, 12, 35, 36] rather than remote eye trackers [37]. Commercial eye trackers currently used in sports are expensive and not suitable to be used actively in many sports situations. So far, there has been very little focus on finding cheap and tailored eye trackers for supporting the athletes with automated tools in their daily training.

Even current eye tracking directed towards sports still requires several new techniques to be developed. Eye trackers should be seamlessly integrated into daily sports training practices and, at the same time, must provide good ergonomy, avoid disturbing the athlete's field of view, be lightweight, comfortable, and robust to changes in the environment conditions. Using eye tracking in sports is likely to require (i) high speed processing to analyze a large amount of eye tracking data in a short time, (ii) high accuracy to identify the athletes' ocular activities, (iii) high robustness to changes in the surrounding light conditions (i.e. athletes may be moving constantly during a match), and (iv) flexible and easy setup of the eye tracker. Solving these open research problems has a broad and meaningful impact in the field of eye tracking research in general.

1.3 Using Priors in Head-Mounted Eye Tracker

The main contribution of this Ph.D. thesis is the use of *priors* which consist of using available information known from the problem at hand to improve the accuracy, precision,

and robustness of eye tracking systems in sports. Priors are available from several distinct sources, such as (i) electronic sensors; (ii) information extracted from eye or scene images; (iii) geometric relationships between both eyes; (iv) location of all components involved in the eye tracking session; (v) fully or partially calibrated setup, among others.

Classical problems of head-mounted eye trackers are usually related to (i) the geometry and location of device components; (ii) simple parametrization of the human visual system; (iii) noise added during some eye tracking stage, among others. A deep knowledge about these problems can assist researchers in creating methods for compensating errors in gaze estimation. For example, (1) if the distance between the user and the observed target is prior known, it is possible to compensate the parallax error using the depth information as priors (see Chapter 4); and (2) if electronic sensors provide information about head movements, it is possible to make up for the head rotations using 3D head angles as priors (see Chapter 5). This Ph.D. thesis presents some approaches of using priors to improve the accuracy, precision, and robustness of head-mounted eye trackers (see Part II) and some assessments using sports experiments as study cases (see Part III).

1.4 Outline of this Ph.D. Thesis

This introduction composes Chapter 1 of this Ph.D. thesis and sets the scene for the presented research project.

Chapter 2 introduces an overview of eye tracking addressed to parameterizations, models, techniques, methods and technical details of eye tracking systems. Chapter 3 presents the literature review of the last ten years, where are analyzed some alternatives of the use of eye tracking in sports, specially head-mounted eye tracker in different sports disciplines. It also discusses how the analysis of eye movements could support novice, intermediate, and expert athletes, and technical body of an elite sports team.

The use of available information known from the problem at hand is the kernel and the main contribution of this Ph.D. thesis. For this reason, each methodological chapter presents an approach of using priors to support the eye tracking system in sports. Each chapter discusses how to collect available information from different data sources and how to use priors to support the improvements in gaze estimation and to compensate the geometric constraints in uncalibrated head-mounted eye trackers.

Chapter 4 presents a mathematical model to compensate the influence of parallax error in head-mounted eye trackers. The parallax error happens due to the spatial offset

between the subject's eye and the scene camera, and it causes a significant offset in gaze estimation when the observed targets are at different depths concerning the calibration plane. This model uses the known distance between the subject and the view target as priors, and estimate the actual subject's gaze in depth with high accuracy.

Chapter 5 presents a mathematical model to compensate the influence of head rotations in the gaze estimation. If the subject keeps his/her gaze on a fixed target and rotates the head around Z-axis (in the right-hand rule), the gaze estimation will include a spatial offset error. This model uses the rotation matrix provided by an orientation sensor as priors and compensates the gaze estimation error through least square regression.

Chapter 6 presents a set of novel eye tracking methods that use priors to enhance the robustness of gaze estimation in challenging environments. Most improvements are related to eye feature detection, which use the geometric relationships between both eyes as priors to validate and to compensate errors in the pupil detection. This chapter also presents two new normalization approaches that use corneal reflections and eye corners as reference points to compensate the head movements in remote eye trackers and the eye cameras slippage in head-mounted eye trackers.

A set experiments assisted in the development of the eye tracking methods using priors proposed in this Ph.D. thesis. Chapter 7 gives an overview of experiments in different sports disciplines to collect and analyze eye tracking data involving novice, intermediate, and expert elite athletes during the daily training session of shooting and kayak.

Chapter 8 is the last chapter, and it discusses the conclusions of this Ph.D. thesis. Instead of restating what has been presented throughout this Ph.D. thesis, it presents a final discussion by offering an all-round of using eye tracking in sports and how to achieve high-accuracy and high-robustness in the eye tracking systems using available information known from the problem at hand.

Chapter

2

Overview of Eye Tracking

Eye tracking is an active multidisciplinary field of research, which has shown considerable progress in the last decades [38]. In practical terms, eye tracking is the process of monitoring eye movements to determine the point of gaze or to analyze motion patterns of an eye relative to the head or the environment [17, 39]. The increase in the processing power of personal computers and the improvements in computer vision techniques are some of the most important factors to the advance in the field of eye tracking [39]. However, current eye tracking techniques still present limitations.

There are two important aspects that the developer should analyze during the development of new eye tracking systems, namely (i) accuracy; and (ii) robustness [40]. Accuracy is a measurement that defines the proximity of gaze estimation to the actual subject's point of regard. On the other hand, robustness is a subjective measurement that is related to the performance of the eye tracking methods, mainly concerning eye feature detection and gaze estimation. The use of eye tracking technology is prevalent in both diagnostic and interactive applications [39, 41].

In the field of human-computer interaction, eye tracking systems can provide continued support for the operation of computer applications. The monitoring of human eye behavior is one of the most efficient interaction mechanism, and the analysis of the user's visual attention will be vital for the next generation of human-computer interfaces [42, 43]. Many scientific research studies are developing computer applications with interaction via eye movements, to make the interaction easier, more intuitive, and more efficient [44], some have proposed using eye tracking as the main input device for computers [45, 46, 47, 48].

Eye tracking is also successful in mobile applications, which the use of wearable eye tracking devices allows the subject to perform daily activities without restrictions during an eye tracking session. For example, eye tracking in sports to monitor the athletes' ocular activities during the execution of sports tasks. Next sections address the parame-

terizations, models, techniques, methods and technical details of the most relevant items destined for the development of eye tracking systems.

2.1 Human Visual System

The analysis of internal and external parameters of human eye structure allows the monitoring of the subject's visual attention. One of the most important internal parameters of the human eye is called *fovea*, i.e. a small region of the *retina* where high-resolution images are projected [49]. The fovea is located around 5° from the *optical axis*, but its exact location is different for each person (i.e. an user-dependent feature). This eye feature determines the relationship between the eye pose (i.e. orientation and position) and the subject's *line of sight*. Gaze estimation methods require the monitoring of eye movements, and these eye parameters need to be tracked all the time during an eye tracking session.

The knowledge of the human eye structure is essential to understand the principle of operation of eye tracking methods. However, the use of a model and the information of the eye parameters is only required by some methods, e.g. model-based eye tracking methods typically estimate the eye pose and the line of sight in three-dimensional models. In general, mathematical models should be generic enough to estimate the gaze of a significant variety of people (e.g. Africans, Asians, Europeans, Latins, men, women, young, elderly, with/without glasses or ophthalmic lenses) [41]. Some important aspects require particular attention in the development of novel ocular-parameters-based eye tracking methods. As human eye structure is user-dependent, the generalization of such parameters tends to decrease the gaze estimation accuracy [40].

Gullstrand-Le Grand Eye Model is widely used in the field of eye tracking to create mathematical models of the *eyeball* [50,51,52,53,54,55]. This model can be used directly to estimate the gaze as well as to perform simulations of eye tracking methods. In summary, it is a simplified mathematical model of the human eye, which assumes that the eyeball has the following features (i) two spheres with distinct sizes; (ii) the eyeball performs a rotation movement around a fixed point; and (iii) there is a small angular difference between the *optical* and *visual* axes for each user [47, 49, 53, 54, 55]. Figure 2 shows a lateral vision of the internal parameters of the human eye.

The optical axis starts at the *anterior pole* and ends at the *posterior pole*, which passes through the centers of the *eyeball* E, *cornea* C and *pupil* P (see Figure 2). The visual axis starts at the *fovea* F and ends at the observed target PoR, which passes through

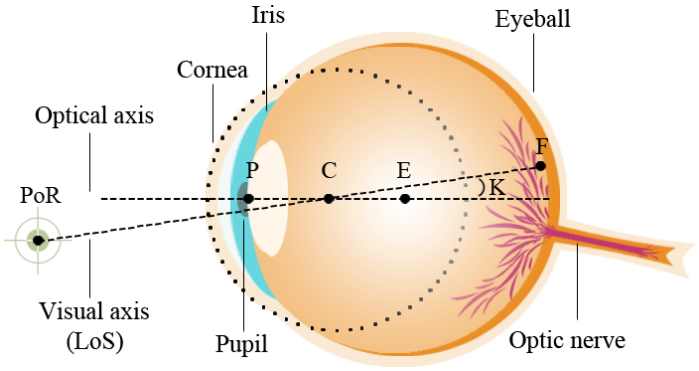


Figure 2: Simplified mathematical model of the human visual system. The eyeball is formed by two spheres, and the *optical* and *visual* axes. The optical axis passes through the centers of the *eyeball* E, *cornea* C and *pupil* P. And the visual axis starts at the *fovea* F and ends at the observed target PoR.

the center of the cornea C (see Figure 2). Therefore, the visual axis defines the highest point of interest in subject's field of view. As the fovea is not localized (exactly) at the posterior pole, the visual axis is slightly deflected from the optical axis. The angle formed by the intersection of the optical and visual axes is called *angle kappa* K, it has an average magnitude of 5° [51, 56] and, in general, has the same magnitude in both eyes [47, 51, 56].

The captured eye images only contain a "virtual" pupil image, because the pupil center in the eye image does not coincide with the actual pupil center in the subject's eye. This phenomenon happens due to the *aqueous humor liquid*. The corneal surface refraction can negatively influence some eye tracking methods, e.g. model-based eye tracking that uses only one camera. The discrepancy between the virtual pupil and the actual pupil increases according to pupil dilation and the angle between eye optical and camera optical axes [49]. Typically, the *refraction index* considered by ocular models used in the several scientific publications is 1.336 [17, 52, 55]. The difference in taking into account the refraction index in the eye tracking method can reflect in a value greater than 1° in gaze estimation [17].

Video-based eye tracking methods analyze eye images that contain only external eye features. One of the simplest approach is to recognize the boundary between the *iris* and the *sclera*, called *limbus* [38, 46, 57, 58]. The limbus does not subject to the corneal surface refraction. However, limbus-based eye tracking methods are influenced by the upper *eyelids* that cover the iris partially and decrease the vertical precision of gaze estimation [38, 46]. Another traditional video-based eye tracking approach analyzes the pupil circumference in eye images. However, it is necessary to use an alternative illumination source to increase the contrast between the pupil and the iris, mainly, in people with dark eyes. Although the eye occlusions are less common over the pupil region, the effects of corneal surface refraction influence the pupil analysis.

2.2 Eye Trackers

An eye tracker is a device responsible for analyzing eye movements and estimating the gaze [39, 59]. Eye trackers have components of different natures, such as mechanical, electronic and optical [60]. These devices can be classified according to the component that plays the most important role, namely (i) *electro-oculography*; (ii) *scleral contact lens/search coil*; (iii) *video-oculography*; and (iv) *video-based combined pupil/corneal reflection* [39]. The hardware components are essential to define the gaze estimation accuracy and eye tracking robustness. In video-based eye trackers, the camera is the most important component once eye tracking begins by capturing images from the eyes.

2.2.1 Eye Tracker Setup

According to some authors (e.g. [51,61]), eye tracker setups can be classified as (i) *remote eye trackers*, if their hardware components do not require to be attached to the subject's body; and (ii) *head-mounted eye trackers*, if their hardware components are fixed on a frame to be worn on the subject's head. Figure 3 shows two examples of vision-based eye tracker setups built with off-the-shelf hardware components.

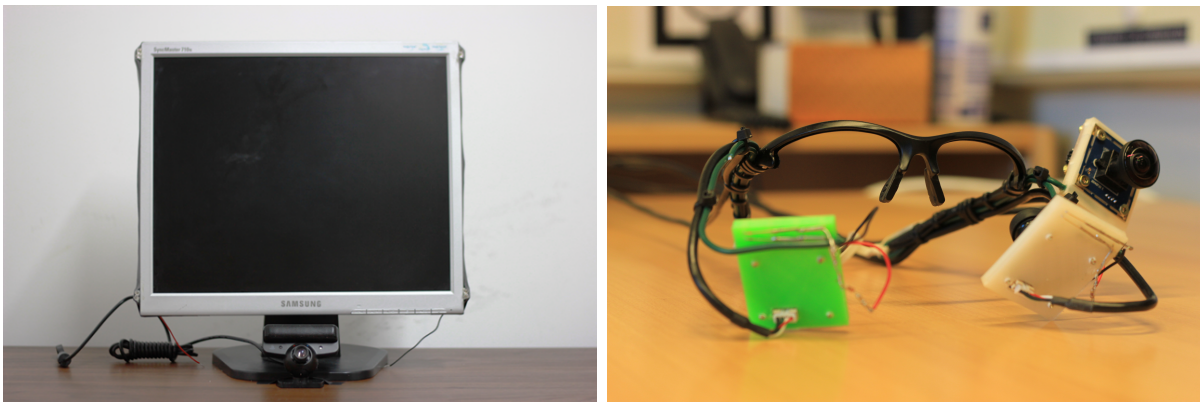


Figure 3: Video-based eye tracker setups built with off-the-shelf hardware. (*left*) a remote eye tracker built with a PlayStation® Eye Camera and four infrared light sources in the screen corners [62]; and (*right*) a head-mounted eye tracker built with two eye cameras and one scene camera in an eyeglass frame.

In general, remote eye trackers are integrated with computer screens to monitor the user's eyes movements from a certain distance. On the other hand, head-mounted eye trackers are usually built on eyeglass frames, caps or helmets, and they are designed to monitor the user's eyes movements from a very close range. Head-mounted eye trackers usually have 2 (*monocular*) or 3 (*binocular*) cameras attached in their physical structure. Eye cameras capture images used in the eye features detection, as shown in Figure 4 (*left*).

On the other hand, scene cameras are used to capture images from the environment which represents the user's field of view, as shown in Figure 4 (*right*).

Figure 4: An example of synchronized images captured by a head-mounted eye tracker. (*left*) the user's eye captured by the eye camera; and (*right*) the scene image that represents the user's field of view.

The use of head-mounted eye trackers is still normally restricted to research centers and specific economic sectors. A widespread use is not yet possible mostly due to high costs associated with commercial eye trackers [63, 64]. Figure 5 shows two different head-mounted eye trackers built for this Ph.D. thesis. These devices cost around $500, and they are cheaper than the commercial head-mounted eye trackers. For example, the Tobii Pro Glasses 2 costs around $30.000 (quotation in September 2016).

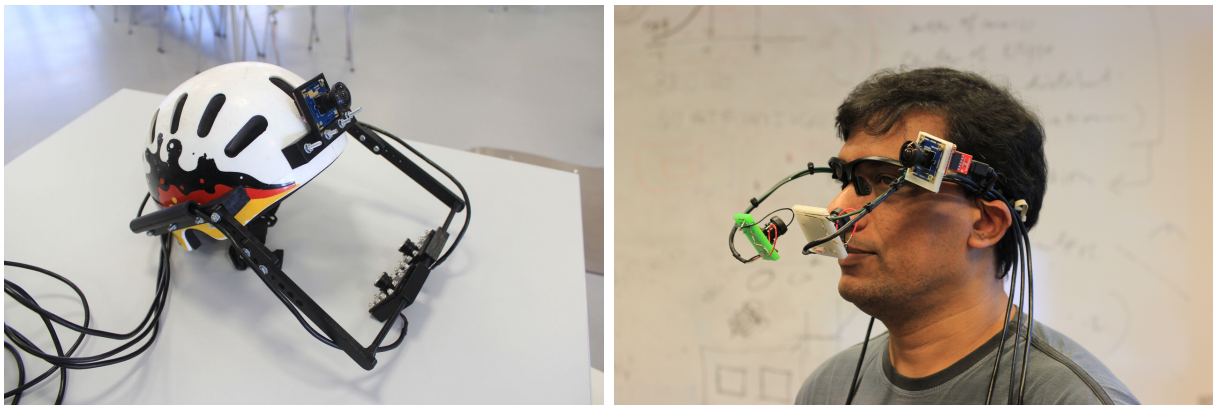


Figure 5: Two head-mounted eye trackers used in sports experiments built with off-the-shelf hardware components. (*left*) a binocular head-mounted eye tracker built in a kayak helmet; and (*right*) a binocular head-mounted eye tracker built in a shooting protection eyeglass frame.

2.2.2 Eye Tracker Calibration

Eye trackers can be classified according to the type of calibration to their hardware components, namely (i) *uncalibrated*, when the parameters and geometry of all eye tracker

components are unknown; (ii) *partially calibrated*, when the parameters or geometry of some eye tracker components (e.g. unknown infrared light source location, but calibrated cameras) are previously known; and (iii) *fully calibrated*, when the parameters and geometry of all eye tracker components are known a prior [17, 65].

Uncalibrated eye trackers usually have the following features (i) head movements are limited for remote eye trackers; (ii) gaze estimation is sensitive to camera slippage in head-mounted eye trackers; and (iii) there is no kind of calibration for any eye tracker components [60]. In this case, it is necessary to perform a new personal calibration at each eye tracking session. The personal calibration is not a complex process since the subjects just needs to look at a set of predetermined targets in their field of view. Despite these limitations, it is possible to achieve similar results (i.e. in terms accuracy and performance) to those of fully calibrated eye trackers [52, 54, 65].

Partially calibrated eye trackers present better gaze estimation accuracy and robustness to head movements in remote eye trackers when multiple infrared light sources are used, and the screen's position is known [17]. Partially and uncalibrated eye trackers are more suitable for the development of eye tracking systems for scientific applications. However, these kinds of calibrations may lead to a more difficult modeling problem [17], mainly the mathematical model used to parameterize the eye structure. According to Hansen and Ji [17], future eye tracking research studies may reveal the potential of partially calibrated setups as well as present more robust approaches to uncalibrated setups.

Some advantages of fully calibrated setups: (i) they are more robust to head movements in remote eye trackers; (ii) the personal calibration is shorter and easier because it requires fewer calibration targets; and (iii) they provide better gaze estimation accuracy. On the other hand, non-commercial fully calibrated eye trackers require a new device calibration every time there is a change in any hardware component [65]. The personal calibration is usually performed only once for each user, with the aim to estimate the user-dependent eye parameters [17, 65]. It means, fully calibrated eye trackers have a complex and tedious device calibration, but a straightforward and fast personal calibration.

For eye trackers that use model-based eye tracking methods and stereo cameras, the first step is to align correctly the stereo cameras' optical axes. In general, their optical axes intersect at a distance of approximately 65cm from the screen [49, 66]. The calibration of cameras and other components can be performed with a unit framework for the MATLAB, called "*MATLAB Camera Calibration Toolbox*[1]" [48, 49, 51, 67, 68]. The cali-

[1]See more information on http://www.vision.caltech.edu/bouguetj/calib_doc/

bration can also be performed through OpenCV library[2], which provides a set of classes for manipulating, managing and capturing synchronized stereo videos [66, 67].

2.2.3 Types of Illumination

The accuracy and robustness of an eye tracker are usually related to the quality of its hardware components as well as its underlying mathematical models. Illumination also plays an important role towards the success of an eye tracker. In general, eye trackers can use two kind of illumination, namely (i) *passive illumination*; or (ii) *active illumination* [41].

Passive illumination exclusively uses the environment natural lighting in the entire eye tracking process. The corneal reflections (i.e. *glints*) obtained through passive illumination are not good reference points, because of the inherent lack of control of these illumination sources (i.e. localization, quantity, state – on/off). In general, it is not common to use passive illumination for producing glints on the corneal surface. The use of passive illumination is usual in limbus-based eye tracking methods.

On the other hand, active illumination is often controlled by the eye tracker. It can use different wavelengths, but the use of infrared light sources (IR) is widespread, because (i) they do not distract the user; and (ii) in indoor environments, they do not suffer interference from variations of the environment illumination, which it is possible to maintain a homogeneous illumination condition. The use of active illumination has two purposes (1) to facilitate the detection of eye features; and (2) to produce *glints* on the corneal surface, which can be used as reference points by several eye tracking methods.

Normally, active illumination-based eye trackers use infrared light sources with wavelengths between 780-880nm, because they are close to the visible-light spectrum [41,48,69]. Infrared wavelength is harmless (unlike ultraviolet) to human eyes, and infrared intensity defines the safety level. If an infrared light source has strong intensity, contrary to visible light spectrum, there is no natural defense mechanism in the human eye, e.g. *pupil contraction* to reduce the quantity of light that enters the eye or *reflex* to divert the gaze.

Active light sources used by eye trackers can be classified according to their locations concerning the eye cameras positions, namely (i) *on-axis*; and (ii) *off-axis*. On-axis light sources are placed in the eye camera's optical axis (or very near to it). On the other hand, off-axis light sources are placed in any other location far from the eye camera's optical axis. Some eye tracking methods explore these two types of active illumination to become

[2]See more information on https://goo.gl/TaVAP7

easier the eye feature detection. Figure 6 shows the effects of using both (*left*) on-axis and (*right*) off-axis infrared light sources in a remote eye tracker.

Figure 6: The use of active infrared light sources in remote eye tracker setups to improve the contrast between the pupil and the iris. (*left*) bright pupil effect caused by the use of one on-axis infrared light source; and (*right*) dark pupil effect caused by the use of two off-axis infrared light sources.

To capture light beams in the invisible light spectrum, the eye tracker needs to use night vision cameras, without any filter to block infrared light and with sensors sensible to the infrared wavelength. The use of redundant infrared light sources to produce more than one corneal reflection can improve the robustness of an eye tracking system because it increases the probability of a minimum number of glints (required by some specific eye tracking methods) to be formed on the corneal surface. On the other hand, it is necessary to include an additional process of identifying and correctly labeling all detected glints [70].

2.2.4 Hardware Components and Features

COGAIN Association (*Communication by Gaze Interaction*) maintains an up-to-date catalog of currently available eye trackers[3], categorized into open-source gaze tracking, freeware eye tracking, low-cost eye tracking, among others. In this catalog, it is possible to obtain information about hardware components, features, and requirements to build a low-cost eye tracking system for both remote and head-mounted eye trackers.

The quantity and type of components used to build an eye tracker are important factors to achieve good accuracy, precision, and robustness in the gaze estimation. It is important to note that the eye model used can also influence the gaze estimation directly. In some cases, the accuracy, precision, and robustness of an eye tracker can be improved only by some changes in the mathematical model, without changing any component.

[3]See more information on http://wiki.cogain.info/index.php/Eye_Trackers

2.3 Eye Tracking Methods

The first scientific studies of eye movements were on text reading. The French ophthalmologist Louis Émile Javal studied eye movements through naked-eye observation during text reading sessions by children [71]. Javal concluded that the readers' eyes do not move continuously along the texts, but performed brief pauses (*fixations*) over small regions of interest and fast eye movements (*saccades*) in the direction to new areas. The human eye behavior is still a theme of continued debate and research [72]. Nowadays, researchers know that many other specialized eye movements provide different types of visual information and can be analyzed in distinct aspects and goals. In this sense, the eye-tracking research field offers support to various knowledge areas that use eye information.

Human eye movements provide valuable information that can be used within several distinct research studies. In general, eye tracking applications can be classified as (i) *diagnostic*; and (ii) *interactive* [39]. In diagnostic applications, the collected gaze data are used with quantitative evidence of the subject's visual attention. On the other hand, interactive applications use gaze data for human-machine communication by the subject's eye movements. Due to some aspects such as feasibility, reliability and accuracy, eye tracking technology has received high interest in the conception of new interactive interfaces based on gaze direction [69]. In the field of Computer Science, it common the use of eye tracking as main interaction mechanism for computer systems.

2.3.1 Eye Tracking Classifications

The are different classifications for eye tracking methods in the scientific literature, based on various aspects such as (1) setups or hardware components used by the eye tracker [38, 73, 74]; (2) compositions of the gaze estimation method [38, 41]; (3) mathematical models used to parameterize the human eye structure [51, 54, 68, 75], among others.

Eye tracking methods can be classified according to whether there is physical contact of any eye tracker component with the user's body, namely (i) *intrusive*; and (ii) *non-intrusive* [38, 73, 74]. Some years ago, this classification was used to distinguish between eye trackers based on measurement components (e.g. electrodes, scleral contact lens/search coil) and eye trackers based on computer vision techniques [39].

There is also a classification based on the input data used to gaze estimation methods, namely (i) *feature-based*; and (ii) *appearance-based* [38, 41]. Feature-based eye tracking methods explore the local eye features from processed eye images (e.g. pupil, iris, glints).

These methods are less sensitive to variations in the environment lighting and are more tolerant to the head movements, e.g. using glints as reference points. On the other hand, appearance-based eye tracking methods are trained with a set of sampled eye images that incorporates appearance variations. In general, these methods are more robust to gaze estimation (even using low-resolution images), and more sensitive to head movements [38].

Eye tracking methods can also be classified according to the mathematical model used to parameterize the eye structure and estimate the gaze, namely (i) *interpolation-based*; and (ii) *model-based* [51, 54, 55, 68, 75]. In general, interpolation-based eye tracking methods use a *non-linear second order polynomial* or a *geometric transformation* to map coordinates from the image plane to the screen/scene plane. On the other hand, model-based eye tracking methods use complex mathematical models to parameterize eye features, which their parameters are computed accurately for each user in three-dimensional space.

Both interpolation-based and model-based methods use distinct mathematical models and typically estimate the gaze with high accuracy (i.e. $error < 0.5°$). However, model-based methods are naturally more tolerant to head movements than interpolation-based methods [56]. Furthermore, model-based methods estimate the line of sight in 3D, while interpolation-based methods estimate the point of regard [75]. Another important difference is the personal calibration because interpolation-based eye tracking methods require visualizing several targets in the screen/scene space, and some model-based eye tracking methods require visualizing only one target in the three-dimensional space [51].

In interpolation-based eye tracking methods, a mapping is used to define the relationship between eye features extracted from input images and the gaze estimation. In remote eye trackers, this approach is sensitive to head movements [38, 56, 60, 61]. Thus, an additional normalization approach based on reference points is necessary to compensate head movements, or the user has to hold the head still during the entire eye tracking session. For head-mounted eye tracker, this method works well if the relation between eye camera, scene camera, and user's eyes do not change during the eye tracking session.

In model-based eye tracking methods, information about global coordinates — of eye features and eye tracker components — are used to estimate the line of sight in 3D [39, 47, 56]. It means, the gaze estimation on a target in the three-dimensional space can be obtained through the identification of the intersection point between the line of sight and the observed target. The model-based approach is naturally tolerant to the head movements because the gaze estimation does not assume a user's specific position or depends on calibrated parameters that are optimized for a user's particular position [56].

2.3.2 Video-Based Eye Tracking Methods

The knowledge of the human visual system is a valuable asset to understand the principle of operation of video-based eye tracking methods. The use of a mathematical model and the knowledge of specific eye feature are only required for some video-based eye tracking methods, e.g. model-based methods that typically estimate the eye pose and the line of sight in 3D. Internal and external eye features are essential to understand and to develop video-based eye tracking methods. However, it is possible to estimate the gaze by the analysis of only external eye features (e.g. pupil, glints, iris) available in eye images.

Pupil-Center-Corneal Reflection Methods (PCCR) PCCR methods are the most traditional video-based eye tracking methods [56]. These methods estimate the gaze by the analysis of the relation between the pupil center and reference points on the corneal surface. In general, these methods require the use of active illumination. Some PCCR-based eye tracking methods use both on-axis and off-axis illumination to detect easier the pupil center. However, it is possible to use only one type of illumination, and even it will be generated glints on the corneal surface. Furthermore, an on-axis illumination is not always convenient for the pupil center segmentation. In most of the eye tracking applications, PCCR methods present good gaze estimation accuracy ($error < 0.5°$) [56].

Iris-Center-Corneal Reflection Methods (ICCR) ICCR methods are among the simplest video-based eye tracking methods. The gaze estimation uses the relation between the center of the iris and corneal reflections [38, 46, 57]. In general, the main disadvantage of these methods are the eyelids that partially cover the iris and decreases the vertical accuracy of the iris detection and gaze estimation [38, 46].

The main differences between ICCR and PCCR methods are the following (i) in ICCR methods, the center of the iris is used instead of the pupil center; and (ii) in ICCR methods, reference points (i.e. glints) on the corneal surface are obtained through visible-light spectrum components (e.g. corneal reflection of the screen) or through active illumination components (e.g. incandescent lamps, flashes). It is still possible to use some additional face/eye feature as reference point [38, 57].

In general, ICCR methods use passive illumination, and they work well for both indoor and outdoor settings. The basic mathematical model of ICCR methods is similar to the model used by PCCR methods. However, in ICCR methods the analyzed vector starts at the center of the iris and ends at a reference point. Figure 7 shows two example of eye images processed with both ICCR and PCCR eye tracking methods.

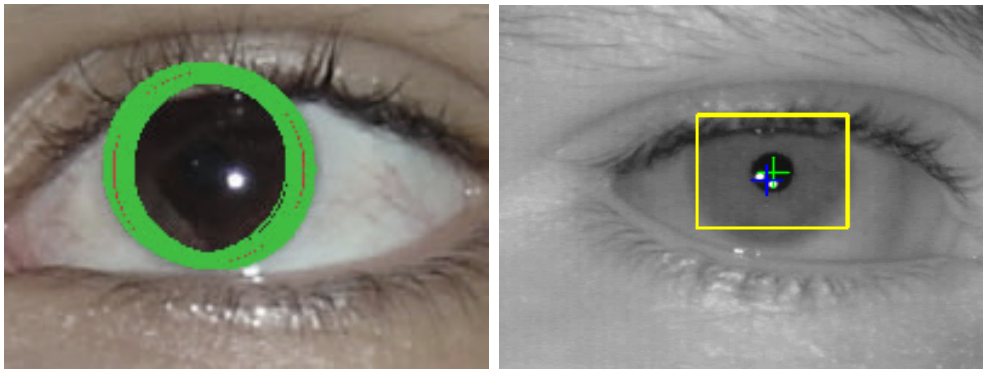


Figure 7: An example of two video-based eye tracking methods that used external eye features to estimate the gaze. (*left*) ICCR methods estimate the center of the iris through the analysis of the boundary between the iris and sclera in visible-light spectrum; and (*right*) PCCR methods usually use active illumination to improve the contrast between the pupil and iris, and to create reference points on the corneal surface.

Cross-Ratio Methods (CR) The original CR method was created to overcome the major limitations of PCCR methods [76]. CR methods require using four off-axis infrared light sources attached to the screen corners to create a polygon of glints on the corneal surface. In this case, the polygon represents a screen projection on the user's cornea, and the pupil center accounts for a forecast of the gaze estimation. From the quadrilateral formed by the corneal reflections and the pupil center, it is possible to use a projective geometry invariant to estimate the user's gaze.

The original CR method has low accuracy, once it is based on a simple mathematical model [77]. Furthermore, this model is sensitive to noise produced by occlusions and presents high errors rates in gaze estimation. In the last years, some improvements have been proposed for the theoretical limitations of CR methods [52, 54, 55, 65, 77, 78, 79]. CR method was designed to tolerate head movements, at the same time, do not require any device or personal calibration [54]. However, in practice, the original CR method is intolerant to large head movements, in special over the Z-axis [54, 55].

Homography-based eye tracking methods are others improvements to the original CR method, which are based on a geometric transformation from the eye plane to the screen/scene plane. It is also possible to normalize a set of corneal reflections in a normalized space to compensate the influence of head movements in remote eye trackers [65]. Homography normalization methods model better the displacement between the optical axis and visual axis compared to the original CR method, and they require at least four calibration targets to create the geometric mapping used to estimate the user's gaze [65].

Model-Based Methods In general, model-based eye tracking methods use a pair of stereo cameras and at least two infrared light sources. These methods have four steps

to identify internal and external eye parameters as illustrated in Figure 2, namely (1) detection of the spherical center of the cornea; (2) estimate of the eye's optical axis; (3) estimate of the eye's visual axis; and (4) estimate of gaze in the three-dimensional space.

Model-based eye tracking methods use the reconstruction property of the three-dimensional triangulation [80] and the convex mirror property [81] to get the global coordinate of the spherical center of the cornea C [56]. For this, it is necessary to use at least two infrared light sources to generate two glints on the corneal surface [56]. In this case, the light beams emitted by infrared light sources form two vectors that intersect themselves exactly in the center of the cornea C.

It is necessary to get the global coordinate of the pupil center to estimate the eye's optical axis. This coordinate can be estimated through reconstruction by three-dimensional triangulation from the extracted information of the stereo images. Remember that, due to the refraction caused by the liquid of the aqueous humor, it is possible to view only the virtual pupil. However, according to the refraction law of the convex mirrors, even with the shift of virtual pupil in the eye images, its center (P') remains in the eye's optical axis [56]. Therefore, the optical axis can be estimated directly through the vector that passes through the centers of the cornea (C) and virtual pupil (P') [47, 56].

The eye's visual axis estimation is not performed directly because the fovea (i.e. the origin of visual axis) is an invisible eye feature in the context of analyzed eye images. However, as the magnitude of angle kappa is constant for each subject, it is possible to estimate the eye's visual axis from the previously estimated eye's optical axis. The personal calibration plays the role to measure the magnitude of angle kappa. In a fully or partially calibrated setup, the subject looks at a single calibration target in 3D space. In the end, the global coordinate of the 3D gaze estimation can be obtained by the intersection of the visual axis and the surface of an observed object at the scene.

2.3.3 Personal Calibration

The personal calibration creates the mapping used to estimate the user's gaze. Eye tracking methods need to estimate a set of geometrical parameters through a personal calibration [17]. These parameters can be calculated before each eye tracking session or just once for each user [56, 65]. The requirements of personal calibration depend strongly on the eye tracking method used by the eye tracker.

For example, some interpolation-based eye tracking methods use a *non-linear second*

order polynomial with 12 unknowns [38,57,61,68,79]. The values of these unknowns (a_i e b_i) can be calculated by *Ordinary Least Squares* (OLS) method, which estimates the best adjustment for this set of data during the personal calibration. Before an eye tracking session, a set of calibration targets are shown to the user by a period of time while the eye tracking system creates the mapping which minimizes the sum of squared residuals between estimated and observed coordinates. Equation 2.1 can be used to estimate the gaze through a non-linear second order polynomial:

$$\begin{aligned} u_x &= a_0 + a_1 v_x + a_2 v_y + a_3 v_x v_y + a_4 v_x^2 + a_5 v_y^2 \\ u_y &= b_0 + b_1 v_x + b_2 v_y + b_3 v_x v_y + b_4 v_x^2 + b_5 v_y^2, \end{aligned} \tag{2.1}$$

which, (u_x, u_y) denotes the coordinates in the screen/scene domain and (v_x, v_y) denotes the coordinates in the image domain. The assumed values for (v_x, v_y) depend on the type of the eye tracking method used to estimate the gaze. For example, in PCCR methods the magnitude of the vector between the pupil center and the reference points are used. If the user views 9 calibration targets during the personal calibration, each target will produce two equations. Thus, the system will use 12 unknowns and 18 equations [38,57,61].

Chapter 3

Overview of Eye Tracking in Sports

When it comes to sports, an important aspect is understanding the role of visual strategy in sports practice [82]. Sports performance analysis is primarily an observational task, and it aims to improve the athletes' performance [83]. However, there are also other judging contexts of sports analysis. With the athletes' eye information added to the kinematics of human movements, environment information, and results, it is possible to identify and analyze athletes' behaviors patterns and thus to improve some of their abilities.

In a global sports market where positive results mean thousands or millions of dollars, advances on performance can be truly significant. Focusing, attention, concentration, consistency, selectivity of perception, and ability to coordinate many actions at the same time [84] are examples of improvements in cognitive areas that can be achieved by athletes' behavior analysis. Some investigations use exclusively observational methods, while others use computerized notational analysis [83]. Nowadays, a common method to get different data for sports analysis is tracking the athletes in many ways, e.g. using an eye tracking system to obtain the athletes' eye information.

Sports analysis is a short way to tell performance analysis of sports. According to O'Donoghue [85], performance analysis of sports is the investigation of actual sports performance or performance in training. The purposes of performance analysis of sports are (i) analysis of technique; (ii) analysis of effectiveness; (iii) tactical analysis; (iv) movement analysis, and (v) analysis of decision making [85]. The sports performance analysis can be done before, during, or after daily training sessions and competitions [85].

According to Passos et. al [86], performance analysis is concerned with the analysis, design, and evaluation of sports systems, which is composed of several layers, as the disciplines physiology, biomechanics, psychology, pedagogy, and sociology. When the primary goal of performance analysis is to understand how the athlete performs related

to decisions, then psychology plays the most import role because the study is focusing on cognitive aspects. The psychology studies have been used eye trackers to investigate how eye movements are related to cognitive processes during different tasks [6]. Eye tracking is also an useful tool to provide detailed performance measures of athletes [8, 9, 10, 11].

3.1 Eye Tracking for Performance Optimization

Investigations regarding gaze behavior use several gaze data measures to provide analysis and to produce the results. Gaze behavior includes gaze direction, gaze estimation, the field of view, quiet eye, visual search, visual fixations, external focus, pupillometry, eye movement amplitude, among others. Combining the collected gaze data with information from the environment, kinematics of human movements, and head movements it is possible to have unlimited varieties of interpretation and analysis of behaviors, causes and effects and other factors, which can influence the athlete's performance results. In this sense, several studies regarding athlete's performance have analyzed various aspects.

3.1.1 Novices versus Experts Athletes and the Use of Time

The performance of the elite athletes depends on the ability to make accurate and timely decisions [87]. According to Fegatelli et al. [7], the novice athlete needs to learn observing all relevant information in a complex and challenging environment using the time efficiently. The best athletes know the best way to collect visual information [7]. To investigate the visual behavior and differences between novice and experts athletes, Hancock and Ste-Marie [88] presented an experiment with ice-hockey players, analyzing the visual search pattern, eye fixations, and decision making. The authors concluded that high-level ice hockey make better decisions than lower-level athletes.

Vansteenkiste et. al. [89] measured the gaze location of the novice, intermediate and experts volleyball players regarding on the strategies used for each group during daily training sessions. They found differences in the athletes' visual search patterns corresponding to their levels of experience. Paeglis et al. [8] presented an analysis of ocular activities from elite junior basketball athletes during throwing training sessions. The evaluation assessed the performance of six junior basketball athletes during two years. After only one year of training, their free throw rates have significantly improved. Paeglis et al. [8] concluded that novice players need more time to make quality decisions before throwing the ball compared to expert basketball players.

Alder et al. [90] investigated the visual search behavior of badminton athletes concerning the responding time to the kinematics of their opponents' movements. The experiment used videos of opponents projected on a life-size screen, and the participants should react to the video as they would play in a real badminton match. The data analysis evaluated the athletes' gaze behavior related to fixations, the number of fixations, fixation duration (minimum 120 ms), and final fixation duration. The authors concluded that expert athletes are more accurate at responding to opponents' actions compared to novices athletes. In summary, novice athletes require more time to fixate on kinematic locations [90], which is the best strategy to defend from the opponents' attacks.

3.1.2 Attention and the Influences of Anxiety

Human gaze behaviors are strongly related to cognitive factors [91]. For example, studies show that the effect of anxiety in athletes on both cognitive and motor performance is an aversive emotional state that occurs as a result of threats and pressures [4, 5]. Studies on visual attention are usually associated with the athletes' performance [92], which can be evaluated from different points of views, such as (i) focus of attention, the act of shifting the eyes to information resources or objects of interest [93]; (ii) visual attention, the capacity to focus only on what is relevant to whatever the subject is doing [92]; (iii) attention control theory, predictions about effects of anxiety on susceptibility to distraction, dual-task performance, and task-switching performance [5, 94], among others.

Related to the focus of attention, Hüttermann et al. [10] presented a study to investigate the ability of athletes to devote attention simultaneously to multiple objects into their field of view. For example, when athletes need to focus on a target and on the opponent at the same time, they have two options (1) focus on one object and attend to the other one in the peripheral vision; or (2) focus between both objects and attend to them in the peripheral vision. Hüttermann et al. [10] showed that athletes present better attention performance when focusing attention simultaneously on two stimuli.

Willians et. al [3] tested the predictions of attentional control theory to examine the effect of anxiety on attention control and its influence on performance (effectiveness and efficiency). The experiments conducted with elite level shooters analyzed the quiet eye duration and onset. As result of tests, the authors reported that changes in performance cause a decrease in shooting accuracy. They also noted efficiency changes, as inferred from the reductions in quiet eye durations, more variable and less efficient gun motions, and increase in mental effort invested in the task.

Several studies investigated the focus of attention of athletes with different data analyzing approaches, like observing the gaze estimation of golf athletes [11], the gaze direction of baseball players [95], visual field of soccer coach [96], and the visual fixations of ice-hockey athletes [88], golf players [97], sailors [92], volleyball players [98] and gymnasts athletes performing high bar [99]. Other studies explore aspects as visual attention [100], the external focus of attention [92], attention control [5, 101], attentional effort [14], and predictive or prospective control [15].

3.1.3 Visual Search Behavior and the Quiet Eye

According to Spitz et al. [102], visual search behavior describes a set of eye movements and fixations available in the cognitive system by visual input. According with Neumann and Sanders [103], visual search behavior is a junction of structural level responsible by the information processed in the visual search that defines a visual search pattern.

Some studies regarding measure attention demonstrated a relationship between the efficient attention in aiming tasks and longer quiet eye durations [82]. The quiet eye is defined as the tracking gaze or final fixation that is located in a specific object or position in the field of view within up to 3° of visual angle for a minimum of 100 ms [104]. According Vickers [104], the relations between the time of fixation, athlete's actions and the results, various interpretation can be made regarding attention, distraction, intended to improve the performance results.

Moran et al. [14] developed a study using pupillometry to evaluate attentional effort in quiet eye for equestrian athletes of different levels of ability. The authors monitored the athletes' eyes movements while making critical decisions by watching a video-based show-jumping sequence. The results showed that pupillometry can be used to identify differences in attentional effort related to skill when the quiet eye occurs. Several others experiments have explored the quiet eye, in diverse sports, like in shooting [5], ice hokey [15], golf putting [101, 105, 106], basketball [107], ice hokey [16], among others.

Heinen et al. [99] presented an exploratory study to investigate functional relationships between gaze behavior (specifically visual spotting) and movement kinematics of gymnasts athletes performing high bar. The authors analyzed the gaze direction and the visual fixations of athletes. The results suggested that gymnasts use visual spotting during preparatory giant swings and dismounts on the high bar. There are functional relations between different fixations and specific movement intentions.

3.2 Eye Tracking Hardware and Systems in Sports

Multidisciplinary research groups use eye tracking technologies to perform amount analysis in various sports disciplines. Table 3.1 presents a review of the most popular choices regarding eye tracking equipment applied in sports, related to hardware, software and systems used in combination and their limitations.

Table 3.1: Summarization studies related to used eye tracking equipment in several sports disciplines.

#	Sports Disciplines	Eye Trackers Setups	Eye Feature Detection	Analysis Software	Gaze Analysis	Additional Information
01	High bar [99]	An off-the-shelf head-mounted eye tracker building in a modified bicycle helmet, with a wireless camera (50 Hz) and two IR sources	PCCR based on two glints	WinAnalyze 3D movement-analysis system	Gaze behavior (gaze direction, visual fixations) and movement kinematics	N/A
02	Shooting [5]	ASL Mobile Eye II; and two XM2 Digital Video cameras (for kinematics records)	Dark pupil, glints and 9 calibration targets	Gamebreaker Software (frame-by-frame); and SIMI Motion 6 Software	Gaze behavior (quiet eye, onset) and movement kinematics	The gaze estimation accuracy checked periodically
03	Badminton [90]	ASL Mobile Eye Tracker; and Qualisys Pro-Reflex MCO 1000 (motion capture cameras)	Monocular, PCCR, 6 calibration targets, and 25 *fps*	Adobe Premier Editing Software; and Qualysis Motion Capture System	Gaze behavior (visual search, visual fixations) and opponent's movement kinematics	Eye movement video footage analyzed frame-by-frame
04	Volleyball [98]	ASL 3000 Mobile Eye	Monocular, PCCR, and 7 calibration targets	Avidemux for MAC	Gaze behavior (visual search)	Videos analyzed twice frame-by-frame
05	Sailling [92]	ASL Mobile Eye Tracker; and an action camera	N/A	Intraclass Correlation Coeficient (ICC) calculation; and IBM SPSS Statistics	Gaze behavior (external focus of visual attention) and head orientation measures	Coding videos twice frame-by-frame
06	Golf [11]	ASL Mobile Eye-XG Eye Tacking Glasses	Monocular and 4 calibration targets	ASL Analysis Software	Gaze behavior (point-of-regard)	Recalibration performed at every trial
07	Basket-ball [107]	Head-Mounted Arrington Reasearch	Lightweight and 9 calibration targets	View Point Software	Gaze behavior (quiet eye)	N/A
08	Ice Hockey [88]	EyeLink II SR Research Eye Movement Recorder	Binocular	SPSS	Gaze behavior (number of fixation, average fixation duration, visual search pattern) and decision making	No head movements allowed
09	Equestrian [14]	Tobii T60 System	Binocular, bright and dark pupil	Tobii Studio; SuperLab Pro 4.0; and Quiet Eye Solutions	Pupilometry (pupil size) and gaze behavior (quiet eye)	QE analyzed frame-by-frame

continues on the next page.

Table 3.1 – continued from previous page.

#	Sports Disciplines	Eye Trackers	Eye Feature Detection	Analysis Software	Gaze Analysis	Additional Information
10	Ice Hokey [16]	ASL Mobile Eye Tracker; and Panasonic PV-GS200	N/A	VIA Vision-in-Action; Quiet Eye Solutions; and Statistical Analysis System	Gaze behavior (quiet eye)	ASL Glasses fitted under a modified goaltender helmet
11	Golf [97]	Eye Tracker LC	Binocular, PCCR, two infrared light sources and 9 calibration targets	NYAN Software; and SPSS	Gaze behavior (number of fixation, average fixation duration)	The eye tracking system is tolerant to pupil drift and head range variation
12	Baseball [95]	ISCAN Eye Tracker; and 3DM-GX1 Head Tracker	Binocular and PCCR	N/A	Gaze behavior (gaze position, eye movement amplitude) and Head movement amplitude	ISCAN data converted from digital to analog; and visual inspection to discard blinks images
13	Football [96]	Tobii X60 Eye Tracker	N/A	N/A	Novices vs. experts' gaze behavior (central visual field, gaze direction)	Discrete-time Gaussian additionally used to reduce the effects of random fluctuations and involuntary micro eye movements
14	Golf [105]	ASL Mobile Eye Tracker	PCCR, Realtime connected via 10m fire wire cable	EyeVision ASL Recording Software; and Quiet Eye Solutions	Gaze behavior (quiet eye) and movement kinematics	QE analyzed frame-by-frame
15	Ice Hockey [15]	ASL Mobile Eye Tracker	Monocular and PCCR	VIA Vision-in-Action; and Quiet Eye Solutions	Gaze behavior (quiet eye)	N/A
16	Volleyball [89]	ASL 501 Head-Mounted Optics	Monocular and PCCR	Eyenal	Gaze behavior (gaze location)	N/A
17	Football [108]	ASL Mobile Eye Tracker	PCCR	EyeVision ASL Recording Software; Quiet Eye Solutions; and GazeTracker Software	Gaze behavior (quiet eye) and attentional control (visual attention, anxiety, performance)	Area of interest defined frame-by-frame
18	Golf [106]	ASL Mobile Eye Gaze Registration	N/A	Quiet Eye Solutions	Gaze behavior (quiet eye)	Gaze behavior analyzed frame-by-frame
19	Football [100]	ASL Mobile Eye Tracking	N/A	EyeVision ASL Recording Software; and GazeTracker Software	Attentional control (visual attention, anxiety, performance) and gaze behavior (quiet eye)	Off-line analysis, calibration process connected by fire wire cable and coordinating determined frame-by-frame

End of Table 3.1

Several studies presented experiments performed in indoor simulated environments. In this condition, the noise from the environment illumination is controlled. Besides this, according to Afonso et al. [98] there are differences in results of experiments in in-situ conditions compared to the simulated environments.

Afonso et al. [98] performed two experiments with volleyball athletes to analyze the differences in eye fixations based on two distinct conditions, namely (i) film-based versus (ii) in situ data collection. The authors concluded that in situ, players employed longer eye fixations than in the film condition. According to Afonso et al. [98], the results suggest that mechanisms underpinning skills decision-making in sports differ between film-based and in-situ conditions. In this sense, it is important to observe that simulated environments or projections do not represent the real sports situations and it can influence on results and analysis.

Others issues associated to eye tracking systems are related by different research studies, such as (i) the need to perform the personal calibration periodically [5, 11]; and (ii) the video analysis performed manually and frame-by-frame [14, 90, 92, 98, 105, 108]. Any author related any issue concerning head-mounted eye trackers, such as (i) parallax error; (ii) head rotations; (iii) scene camera's field of view; or (iv) illumination problems.

Part II

Robust Eye Tracking Methods in Sports

Chapter

4

Using Priors to Compensate the Parallax Error

Parallax error is a geometric problem due to the projection centers of the scene camera, eye camera, and user's eyeball are not co-axial [18, 20, 22]. This issue causes significant gaze estimation errors (in the scene view) when the observed targets move at different depth locations than during personal calibration. In addition to the inherent inaccuracy of gaze estimation ($error > 0.5°$), the parallax error will effectively make it hard, if not impossible, to analyze gaze data reliably.

The parallax error is a significant problem of eye tracking in sports, which the observed target (e.g. ball, disk, gate, opponent, stone) moves in depth relative to the athlete. So far most of the eye trackers can not account for the parallax error, and therefore research results are often based on human inspection and estimates on the location the point of regard when the target moves in space [72].

This chapter presents the use of depth information as priors to compensate the influence of parallax error exclusively for head-mounted eye trackers. Several different data sources such as visual tracking, stereo, and electronic sensors can provide the depth information. Thus, Section 4.1 describes the parallax error in more detail, and Section 4.2 describes the proposed parallax compensation model.

Section 4.3 presents an evaluation of the proposed compensation model using simulated data, which consistently improves the gaze estimation accuracy when the target is in both calibration plane and different depth planes. On the other hand, Section 4.4 discusses the improvements achieved using real gaze data collected during a laboratory experiment. This chapter intends to show the possibility of estimating the athletes' gaze actively in sports practices and thus overcome some of the problems related to the gaze estimation in multiple depth locations using head-mounted eye trackers.

4.1 Influence of Parallax Error in Gaze Estimation

Most current head-mounted eye trackers can not provide high accuracy gaze estimation ($error < 0.50°$) due to the parallax error. The scene camera is not on the same axis of the subject's eye to avoid occlusions; however, the scene camera placed on the same axis would reduce the influence of parallax error in gaze estimation. For this reason, it is necessary to find out a solution to compensate the parallax error for general head-mounted eye trackers. Figure 8 shows an example of the geometry of parallax error.

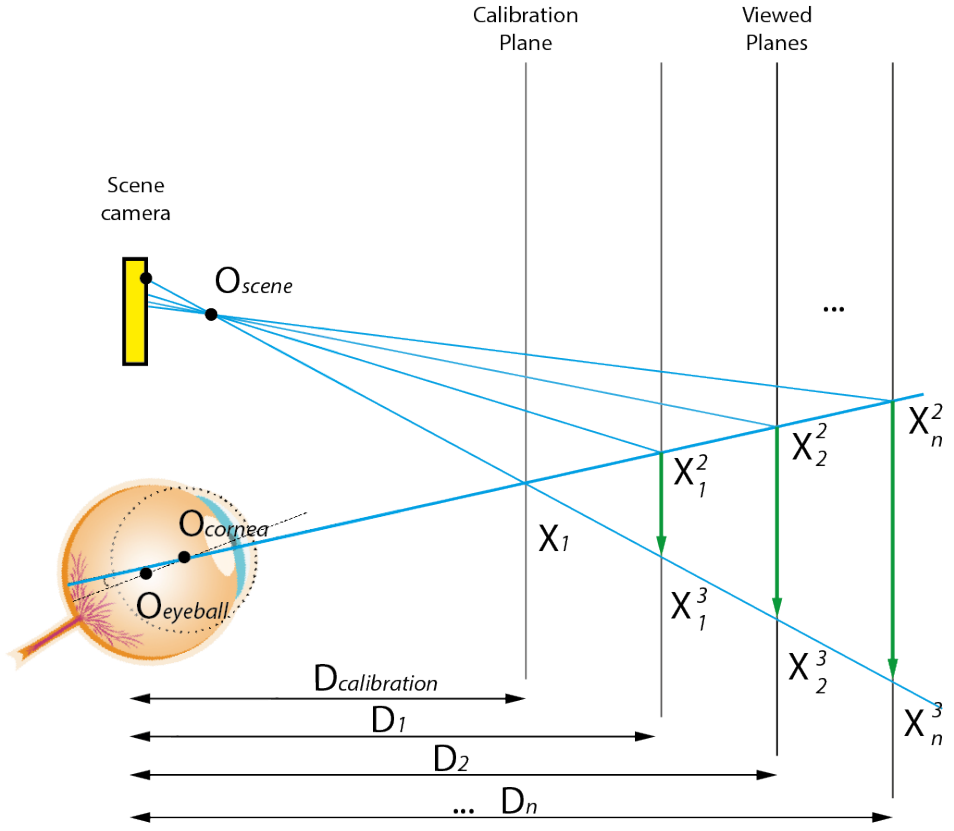


Figure 8: The parallax error geometry in head-mounted eye trackers. In general, the personal calibration of head-mounted eye trackers uses a fixed calibration plane. Targets on calibration plane can be estimated with high accuracy using traditional gaze estimation methods (e.g. cross-ratio, polynomial, homography). However, when the user looks at a target in the same position X_1 but on a different plane D_i, the gaze will be estimated on position X_i^3 instead of the position X_i^2. The green arrows represent the parallax error corresponds to the vector $||X_i^2 X_i^3||$ on the observed plane and a vector $||x^1 x_i^2||$ on the scene plane.

The first parameter to be analyzed is the subject's eye, based on *Gullstrand-Le Grand Eye Model* (see Section 2.1). The second parameter is the scene camera, which is not co-axial with the subject's eye and is based on the pinhole model with a vertical image plane. The last parameter to be analyzed are the planes observed during the eye tracking session. The personal calibration of head-mounted eye trackers uses a calibration plane in a given distance $D_{calibration}$ from the subject. The eye tracking system can provide

high accuracy gaze estimation in the entire calibration plane. The subject's visual axis intersects the calibration plane exactly at the point X^1 and the multiple depth planes D_i at the point X_i^2 $(1 \leq i \leq n)$. When the subject looks directly at the point X^1, the gaze is estimated correctly as the point x^1 on the image plane. On the other hand, when the subject looks at to any depth plane at the point X_i^2, the gaze will be estimated as the same point x^1 instead of the point x_i^2 on the image plane.

According to Mardanbegi and Hansen [19], the parallax error produces an offset vector $||x^1 x_i^2||$ on the image plane, which corresponds to the vector $||X_i^2 X_i^3||$ on the observed plane [19, 20]. Figure 9 shows an example of using an eye tracking system for a curling daily training session. The parallax error appears in different planes on the curling sheet when the athlete look at to the stone (target) far away from the calibration plane.

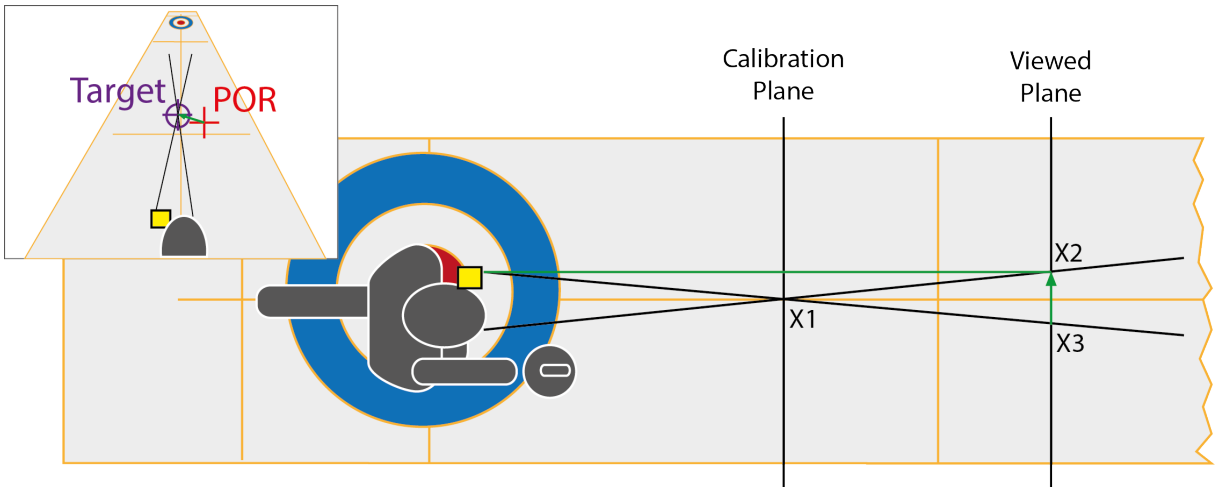


Figure 9: An example of parallax error during a curling daily training session. On the calibration plane, there is no influence of parallax error after the personal calibration. On the other hand, when the stone goes to position X_2 on the observed plane, the gaze will be estimated on position X_3. When the distance between the athlete and the stone is prior known, the proposed parallax compensation model can correct the gaze estimation to the actual position in the scene image, as shown in the small upper picture.

Several research studies proposed solutions to minimize the influence of parallax error in head-mounted eye trackers. For example, Velez and Borah [18] presented an eye tracker that uses hot-mirror glasses in front of the subject's eyes to control the distances and angular relationship between the eye camera, scene camera, and eyes. The hot-mirror is positioned in the eye's optic axis with the aim to reflect images from the eyes and environment toward their respective cameras. This setup removes the parallax error and ensures a wide-angle scene viewing over multiple depth planes. Not all head-mounted eye trackers have an architecture that allows to attach hot-mirror glasses on them, and thus the problem persists. Mardanbegi and Hansen [19] proposed a study to identify the main

sources of parallax error in head-mounted eye trackers. They analyzed the influence of scene camera positions, the calibration and fixation distances on the parallax error. They showed that the angle kappa does not have a significant effect on the parallax error [19,20].

4.2 Parallax Compensation Model

The phenomenon of parallax occurs due to the geometry of head-mounted eye trackers, which is similar to *epipolar geometry* in a stereo vision system [19, 109]. The epipolar geometry is expressed by the point p_{eye} on eye plane and the point p_{scene} on scene camera plane that must lie on a line called *epipolar line*. As shown in Figure 10, if the point P in subject's field of view moves along the line formed by center of the user's eye $O_{eyeball}$ and point p_{eye}, its projection on the eye plane will not change but the projection on the scene plane will change. This movement traces out the epipolar line L_{scene} [109].

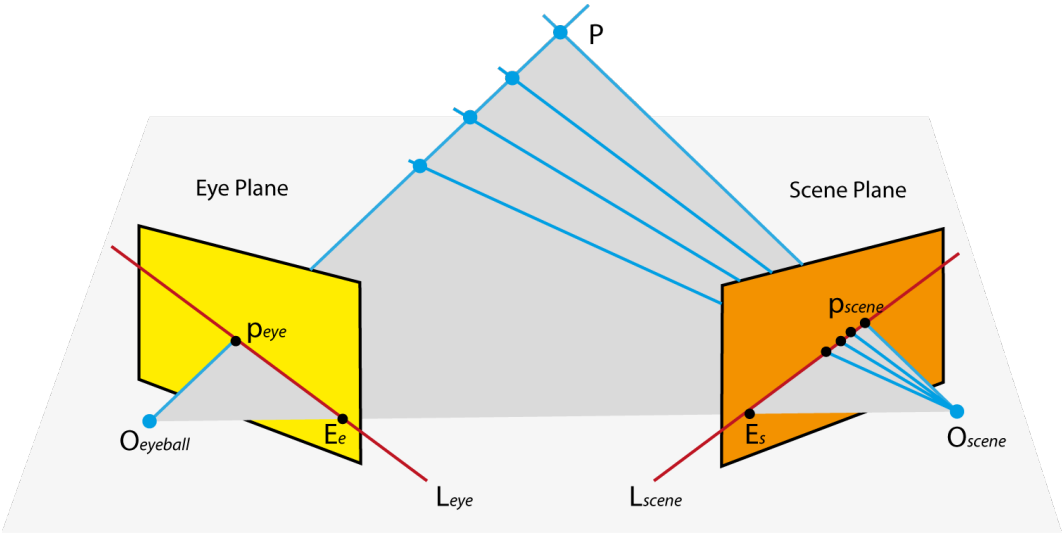


Figure 10: The epilolar geometry of a monocular head-mounted eye tracker.

When the subject focuses on targets at different planes (see Figure 8), the point of regard will move along an epipolar line on the scene plane. Based on epipolar geometry, all epipolar lines intersect at a common point called *epipole*. In this context, the epipole is the point E_s on the scene camera plane. The epipolar geometry uses the *fundamental matrix* F to estimate each epipolar line [110]. To estimate the fundamental matrix, it is necessary at least eight corresponding points in both eye and scene images [111].

The fundamental matrix F encapsulates the intrinsic camera geometry, and it is independent of scene structure [109]. Given F as a 3×3 matrix, Equation 4.1 calculates

the corresponding epipolar line L_{scene} for every point p_{eye} in the eye image plane:

$$L_{scene} = F \times p_{eye}. \tag{4.1}$$

In a binocular head-mounted eye tracker, each eye image with the scene image forms a geometry similar to a stereo vision system. Thus, the binocular eye tracking approach defines two epipolar lines (L_{left} and L_{right}) on the scene camera plane. As the point of regard moves along each epipolar line [19], the intersection between two epipolar lines is very close to the actual subject's gaze. Let the epipolar line from left eye as $L_{left} = ax + c$ and the epipolar line from right eye as $L_{right} = bx + d$. Equation 4.2 presents the proposed binocular gaze estimation method based on the interception of two epipolar lines:

$$g = \left(\frac{d-c}{a-b}, \frac{ad-bc}{a-b} \right). \tag{4.2}$$

Let the scene camera captures images from the environment based on the pinhole model, as shown in Figure 11. The camera center C is the origin of a Euclidean coordinate system (i.e. *center of projection*), which all points in the 3D-space join in C. If Z-axis is the *principal axis* in camera model, then the image is formed in any plane orthogonal to the Z-axis at a distance f (i.e. *focal length*) from the camera center. The principal point p is the intersection between the principal axis and the image plane. The origin of the world coordinate system O is related to C by a rotation R and a translation t.

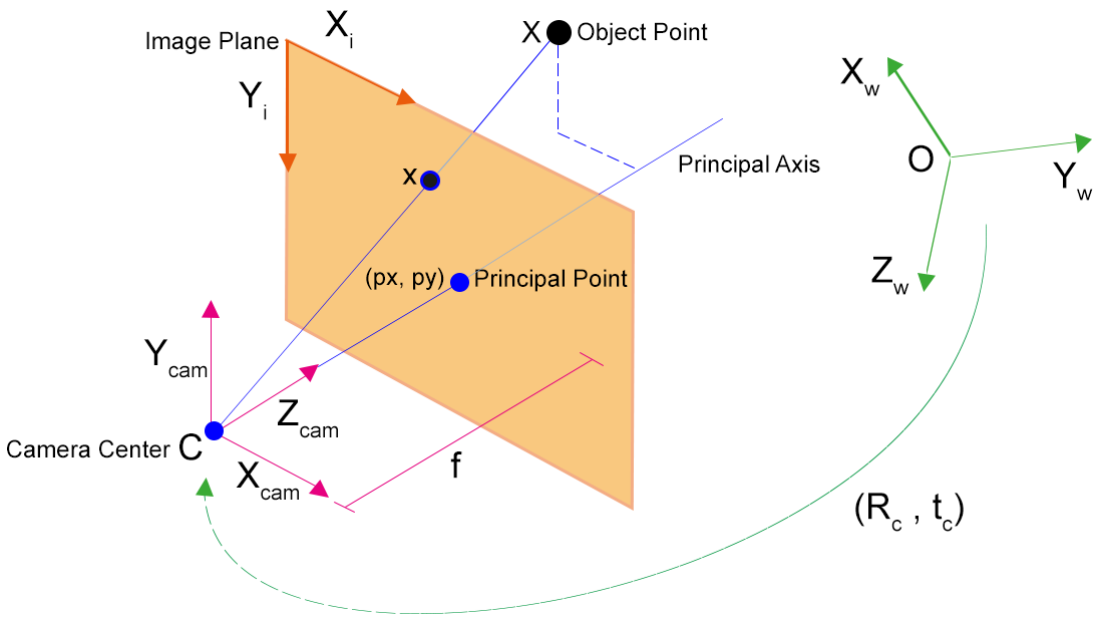


Figure 11: The proposed parallax compensation model is based on the pinhole camera model. This model describes the relationship between the coordinates of a point in three-dimensional space X that moves in depth Z and join in the camera center of projection C.

Let $X = (X, Y, Z)^T$ as a point in three-dimensional Euclidean space, there is a line that join the point X to the center of projection C. This line intersects the *image plane* at point x. Considering the origin of the image coordinate system as the principal point p, the mapping from Euclidean 3D-space $\mathbb{R}^3$ to Euclidean 2D-space $\mathbb{R}^2$ can be described linearly in terms of homogeneous coordinates as Equation 4.3:

$$x \approx \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} X. \tag{4.3}$$

In general, the origin of the image coordinate system is not located at p. The principal point offset is described as (p_x, p_y). There are some non-Euclidean properties of the relationship between the camera coordinate system and the image coordinate system must be accounted for. The captured images usually have some distortion, with respect to aspect ratio α (i.e the ratio of height to width of a pixel) and camera skew s (i.e. skewness in the camera sensor). Thus, Equation 4.4 defines the camera calibration matrix K that describes the intrinsic parameters of a pinhole camera.

$$K = \begin{bmatrix} \alpha f & s & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix}. \tag{4.4}$$

As the world coordinate system is not coincident with the camera coordinate system (i.e $C \neq O$), the camera's location and orientation in the world coordinate system is defined by (R, t). Thus, the relationship between the point X in camera and world coordinate systems depends on a rotation and a translation, as defined in Equation 4.5:

$$X_{cam} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} X, \tag{4.5}$$

which, $t = -RC$. The angles from R and C are known as the extrinsic parameters of a pinhole camera. Both intrinsic and extrinsic camera parameters define the full mapping from Euclidean 3D-space $\mathbb{R}^3$ to Euclidean 2D-space $\mathbb{R}^2$ of a pinhole camera. Equation 4.6 defines the pinhole camera projection matrix P, which consists of Equations 4.4 and 4.5:

$$P = K[R|t]. \tag{4.6}$$

Even proposed gaze estimation method (see Equation 4.2) still suffers influence of

parallax error. For this reason, it is proposed a novel parallax compensation model based on pure translation motion, i.e. a planar motion case where there is no rotation [110,112].

After estimating the gaze, it is necessary to correct the parallax error based on the depth information between the observed target and the calibration plane. As the cameras are stationary in the head-mounted eye tracker, the proposed parallax compensation model assumes that only the targets undergo a translation $-t$ and there is no rotation $R = 0$. In this case, the 3D points move on straight lines parallel to the direction of t. One may assume that the calibration plane and observed plane are respectively $P_{calibration} = K[I|0]$ and $P_{observed} = K'[I|t]$. Equation 4.7 calculates the fundamental matrix when there is no rotation ($R = I$) and both camera matrices are the same ($K = K'$):

$$F = [e']_{\times} K K^{-1} = [e']_{\times}, \tag{4.7}$$

which the notation $[e']_{\times}$ is a rank 2 skew-symmetric 3×3 matrix. For example, if the target plane translation is parallel to the calibration plane Z-axis, then $e' = (0, 0, 1)^T$. In this case, Equation 4.8 defines the fundamental matrix for translation on Z-axis:

$$F = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{4.8}$$

If the gaze estimation g in the calibration plane is normalized as $g = (x, y, 1)^T$, then from $g = PG = K[I|0]G$, which the coordinates of point G in Euclidean 3D-space $\mathbb{R}^3$ are $(X, Y, Z)^T = ZK^{-1}g$, where Z is the depth of the point G from the observed plane along the principal axis of the calibration plane. It follows from $g' = P'G = K[I|t]G$, Equation 4.9 estimates the user's gaze g' without the influence of parallax error:

$$g' = g + Kt/Z, \tag{4.9}$$

which depends on the magnitude of translation t and the inverse depth Z [110]. The proposed model uses the *distance* Z between the observed target plane and the calibration plane as *priors* to estimate the actual subject's gaze g'. The projection matrix of calibration plane $P_{calibration}$ represents the origin of pure translation motion as defined in Equation 4.10:

$$P_{calibration} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{4.10}$$

which, there is no information about translation or rotation. In turn, the projection matrix of observed target plane $P_{observed}$ is estimated from the fundamental matrix F that defines the three-dimensional axis which the pure translation will be performed (e.g. Equation 4.8). The camera matrix K, translation matrix t, and rotation matrix R (that is always equals to identity matrix I) are extracted from the difference between $P_{calibration}$ and $P_{observed}$. Finally, the gaze estimation without the influence of parallax error is calculated using Equation 4.9, which Z is the *prior information* related to the distance in meters between the observed target plane and the calibration plane in the real world.

4.3 Assessment using Simulated Gaze Data

This section presents the evaluation of the proposed gaze estimation method (based on binocular eye tracking systems and epipolar geometry, see Equation 4.2) and the proposed parallax compensation model (based on pure translation motion and depth distance as priors, see Equation 4.9) using simulated gaze data from an entirely controlled simulation environment. The used MATLAB eye tracker simulator allows controlling all eye tracker settings as well as all human ocular parameters [113].

The evaluation using simulated gaze data contains two distinct stages, namely (i) evaluation of proposed gaze estimation method according to when the user observes targets only on the calibration plane (Subsection 4.3.2); and (ii) evaluation of proposed parallax compensation model according to when the user observes targets on the multiple depth planes (Subsection 4.3.3). The evaluation investigated the error distribution in gaze estimation concerning (1) refractive index of aqueous humor [α]; (2) number of targets in personal calibration [N]; (3) horizontal [γ] and vertical [β] angle offset between optical and visual axes [aka *angle kappa*]; (4) noise in eye feature estimation [$Pc + \lambda$]; and (5) targets moving in depth along to Z-axis with respect to the calibration plane.

4.3.1 Apparatus

The MATLAB eye tracker simulator [113] simulated a binocular head-mounted eye tracker with two eye cameras (one for each user's eye) and one scene camera (to monitor the simulated environment). Both eye cameras were on the calibration plane in the bottom. The left eye camera was 20 cm to the left from the center of calibration plane, and the right eye camera was 20 cm to the right. In turn, the scene camera was slightly closer to the user position in the simulation environment, between both user's eyes and aligned

with the center of calibration plane (with respect to X-axis). The calibration plane was 55 cm from the user, and the personal calibration used $N = 8$ targets to calculate the fundamental matrices (F_{left} and F_{right}) used to estimated the simulated gaze data.

This evaluation used two distinct eye models with different angle kappa offsets, based on the assessment protocol of Hansen et al. [65]. The eye model E_0 $[\beta = \gamma = 0°]$ is a physically infeasible setup, used only to avoid some particular eye biases in gaze estimation. On the other hand, the eye model E_1 $[\beta = 1.5°, \gamma = 4.5°]$ contains a more realistic angle kappa for the human visual system. During each test (see full description in introduction of Section 4.3), the proposed gaze estimation method generated the gaze error distribution of 4.096 observed targets distributed in a 64×64 matrix on the observed plane.

4.3.2 Evaluation of Proposed Gaze Estimation Method

Refractive Index of Aqueous Humor This test evaluated the influence of the refraction index in proposed gaze estimation method. According to Hansen and Ji [17], the refractive index of aqueous humor has a constant value around 1.336. Table 4.2 presents the influence of the refractive index of aqueous humor based on two eye models (E_0 and E_1) and with the presence or not of refraction. The results show that angle kappa offset presents a notable difference in proposed gaze estimation method when comparing the gaze data from E_0 and E_1 eye models. On the other hand, there is no influence of the refractive index because the gaze error distributions do not change for each eye model.

Table 4.2: The influence of the refractive index of aqueous humors [1.336] in proposed gaze estimation method, using two eye models, i.e. E_0 $[\beta = \gamma = 0°]$ and E_1 $[\beta = 1.5°, \gamma = 4.5°]$.

Model	Refraction	Maximum Error	Mean Error
E_0	No	$0.23° \times 10^{-4}$	$0.37° \times 10^{-5}$
E_0	Yes	$0.23° \times 10^{-4}$	$0.37° \times 10^{-5}$
E_1	No	$0.79° \times 10^{-1}$	$0.89° \times 10^{-2}$
E_1	Yes	$0.79° \times 10^{-1}$	$0.89° \times 10^{-2}$

Number of Calibration Targets This test evaluated how the number of calibration targets ($8 \leq N \leq 25$) influences the accuracy of the proposed gaze estimation method. Figures 12 and 13 show the gaze estimation accuracy as a function of the number of calibration targets. For both eye models (E_0 and E_1) the minimum number of calibration targets ($N = 8$) presented lowest gaze error distribution. Figure 12 shows that the gaze error is not steady for the eye model E_0. On the other hand, Figure 13 shows a more

stable gaze error distribution for the eye model E_1. The more realistic eye model E_1 did not improve the accuracy of proposed gaze estimation using more than the minimum number of calibration targets required by the personal calibration, in opposite to classical gaze estimation methods [65, 78, 114].

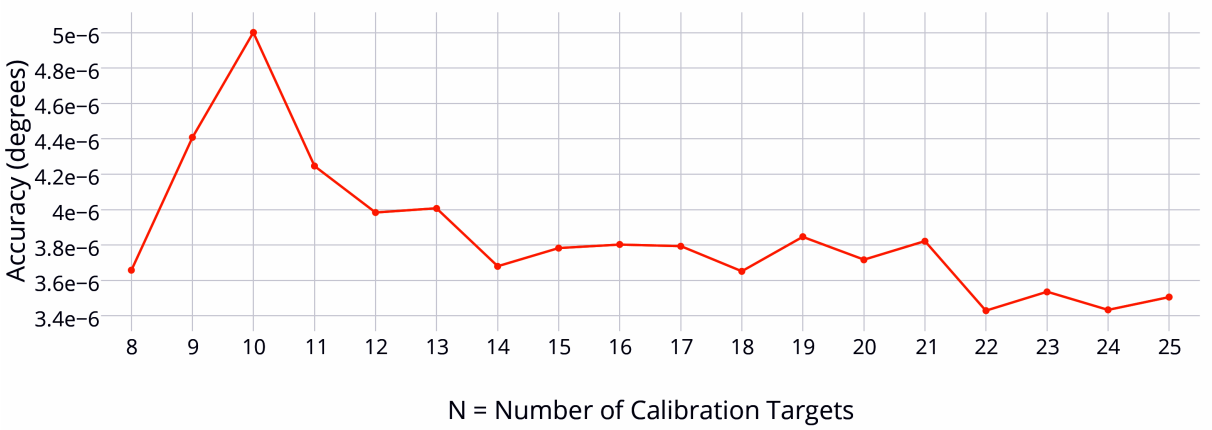


Figure 12: The influence of the number of calibration targets N to the proposed gaze estimation method using eye model E_0 $[\beta = \gamma = 0°]$.

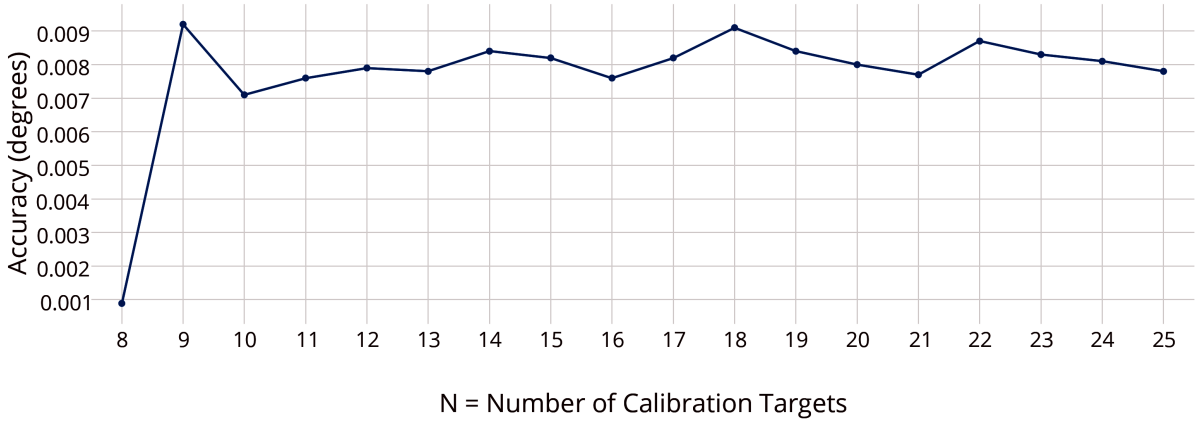


Figure 13: The influence of the number of calibration targets N to the proposed gaze estimation method using a more realistic eye model E_1 $[\beta = 1.5°, \gamma = 4.5°]$.

Angle Kappa Offset This test showed that differences in angle kappa have a significant influence on the accuracy of proposed gaze estimation method. Figure 14 illustrates the influence of different angle kappa offsets within a range of angular horizontal offsets $(-4.5° \leq \gamma \leq 4.5°)$ and a fixed angular vertical offset $(\beta = 0°)$. The results show that the accuracy of proposed gaze estimation method linearly decreases according to angle kappa, i.e. the bigger the angular difference among visual and optical axes (an user-dependent parameter) the lower will be the gaze estimation accuracy. In humans, angle kappa has the same magnitude in both eyes, and its average angle is around 5° [56]. It means, the

proposed gaze estimation method based on epipolar geometry has better accuracy than classical gaze estimation methods, i.e. in the worst case the error is around 0.01°.

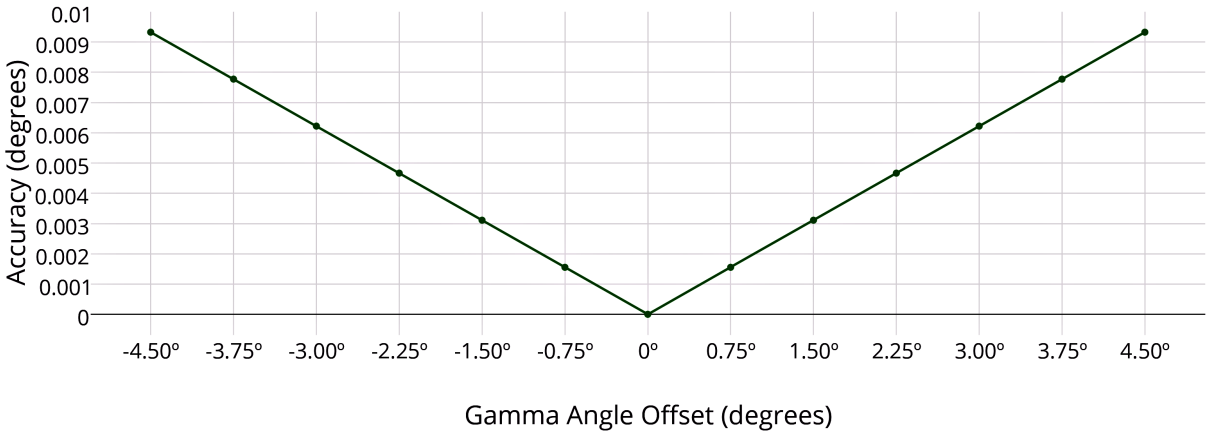


Figure 14: The influence of angle kappa offset to the proposed gaze estimation method. Angle kappa has two angles offsets, i.e. horizontal (γ) and vertical (β). The are results of the influence of angle kappa with $-4.5° \leq \gamma \leq 4.5°$ and $\beta = 0°$.

Noise This test investigated if the proposed gaze estimation method would have the same accuracy in a real eye tracking scenarios, i.e. the robustness of proposed method adding noise to the eye feature detector. The evaluation added controlled noise (λ) in the pupil center Pc before calculating the epipolar line, i.e. $L_{left} = F_{left} \times (Pc_{right} + \lambda)$ and $L_{right} = F_{right} \times (Pc_{left} + \lambda)$. Figure 15 shows a two-dimensional view of the noises in a range of horizontal coordinates and a fixed vertical coordinate, i.e. $Pc = (x + \lambda_x, y)$.

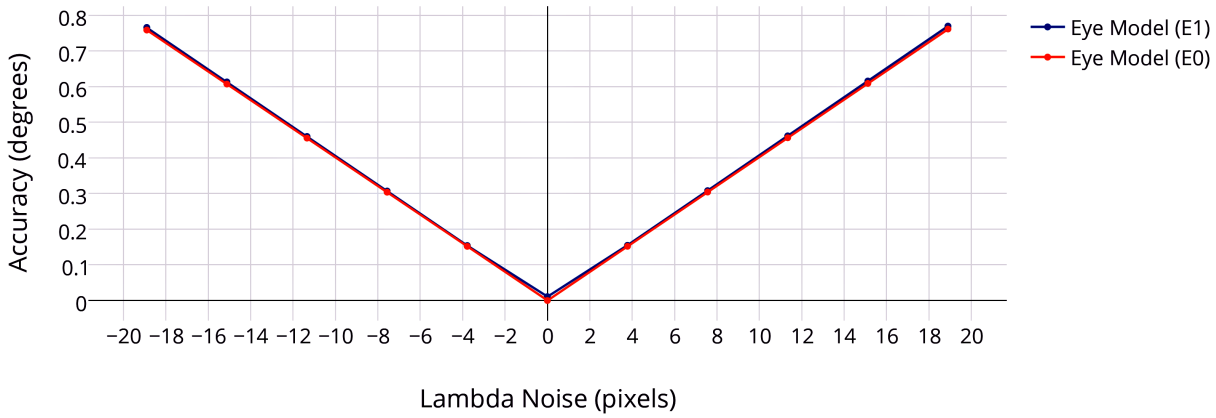


Figure 15: The influence of noise in the proposed gaze estimation method when adding noise to the pupil center coordinate using (red) eye model E_0 $[\beta = \gamma = 0°]$ and (blue) eye model E_1 $[\beta = 1.5°, \gamma = 4.5°]$. The noise ($\lambda$) added to pupil center $Pc(x, y)$ has the following range $-18.90 \leq \lambda \leq 18.90$ in pixels.

The results show that the proposed gaze estimation method only achieves better accuracy than traditional gaze estimation methods when there is no noise. For a real

eye tracking scenario, the proposed gaze estimation method presents an average gaze error around 0.50°. Figures 16 and 17 illustrate the influence of noise in the eye feature detection respectively for eye models E_0 and E_1 in proposed gaze estimation method.

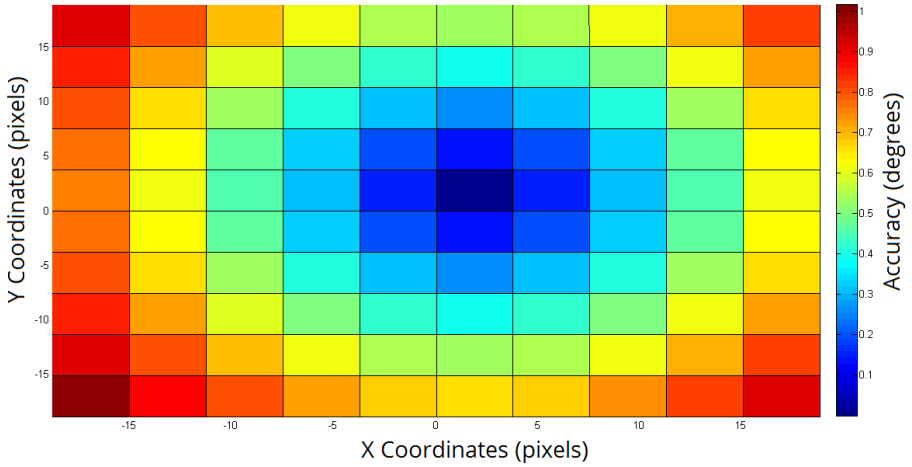


Figure 16: The influence of feature detection noise to the proposed gaze estimation method using eye model E_0 [$\beta = \gamma = 0°$]. The pupil center $Pc(x, y)$ received an spatial noise in the following range $-18.90 \leq \lambda \leq 18.90$ pixels, i.e. ± 5 mm in the simulated environment.

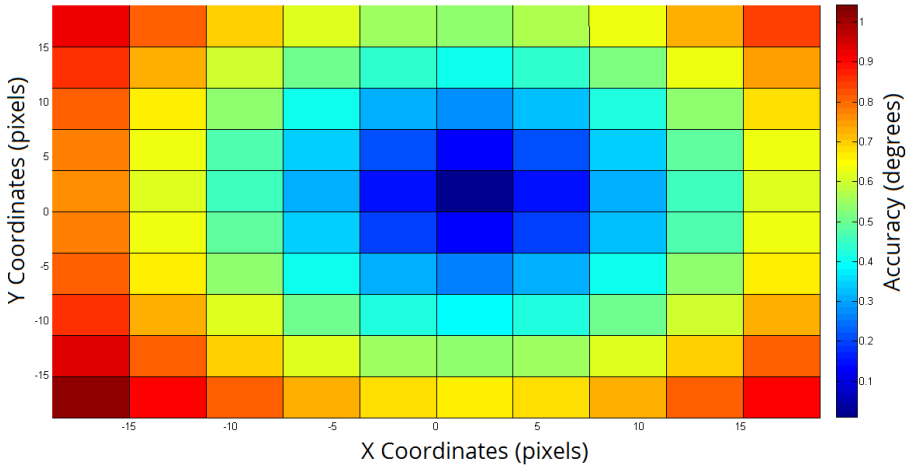


Figure 17: The influence of feature detection noise to the proposed gaze estimation method using eye model E_1 [$\beta = 1.5°, \gamma = 4.5°$]. The pupil center $Pc(x, y)$ received an spatial noise in the following range $-18.90 \leq \lambda \leq 18.90$ pixels, i.e. ± 5 mm in the simulated environment.

For both eye models, the proposed gaze estimation method is sensitive to noise in the eye feature detection. According to Luong and Faugeras [115], fundamental matrix based on the eight point algorithm [111] is very sensitive to noise. The gaze estimation error was small only when $\lambda = 0$. In this case, eye model E_0 presented a gaze estimation accuracy around $0.39° \times 10^{-5}$ and eye model E_1 was around $0.11° \times 10^{-1}$. For $\lambda > 0$ the gaze estimation accuracy is similar for both E_0 and E_1, i.e. the difference mean is $\pm 0.01°$.

4.3.3 Evaluation of Proposed Parallax Compensation Model

The last test with simulated gaze data aimed to evaluate the proposed parallax compensation model. During this test, the simulated eye tracker and user kept still while the targets moved along to Z-axis. The calibration plane was at 55 cm from the user, and the observed targets moved in depth in a range of 35-125 cm from the user in steps of 10 cm. Figure 18 shows the influence of parallax error to the proposed gaze estimation method in a range of 35-125 cm in depth. Figure 19 shows the gaze error distribution when the observed targets were 95 cm from the user. The results showed that for each meter from the calibration plane, the parallax error adds a gaze error around ±2.30°.

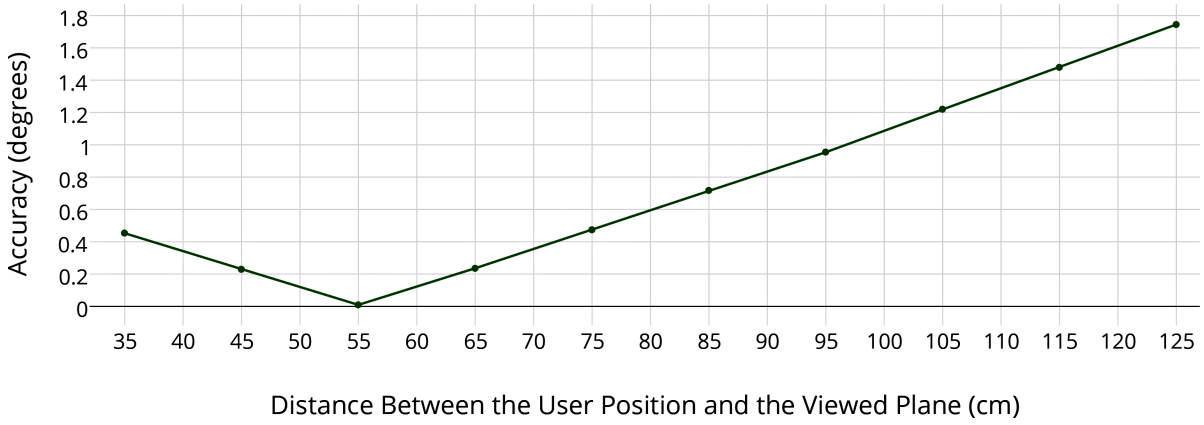


Figure 18: The influence of parallax error to the proposed gaze estimation method. The observed targets moved to 10 different distances from the user position, i.e. in a range 35-125 cm in steps of 10 cm. The results show that the gaze estimation accuracy decrease a lot because the parallax error.

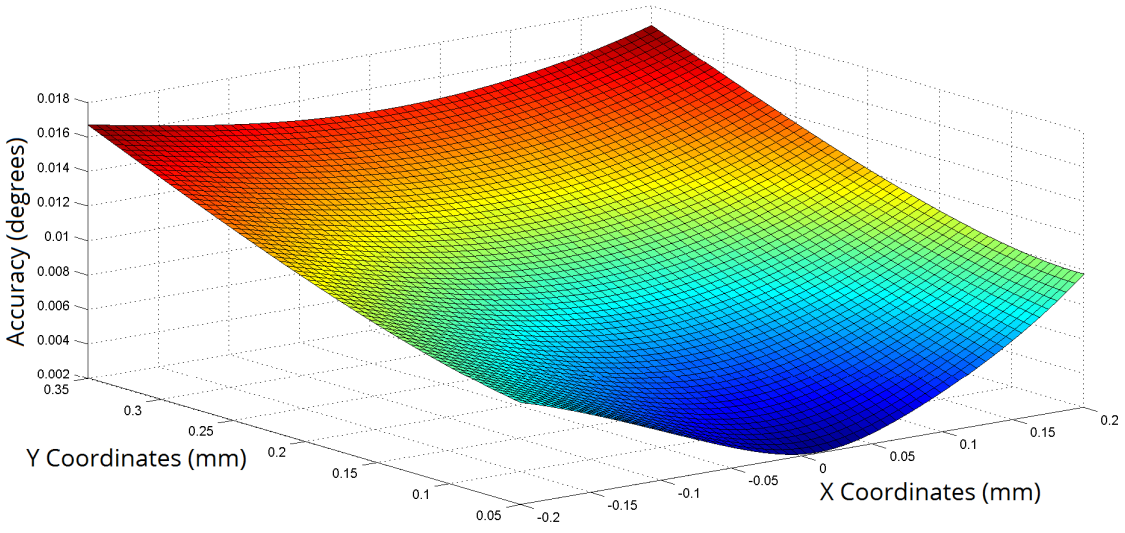


Figure 19: The influence of parallax to the proposed gaze estimation method. After personal calibration, the observed targets moved to 95 cm from the user, and the gaze estimation accuracy was 0.95°.

Figure 20 shows the gaze estimation after reducing the influence of parallax error using the proposed compensation model, in a range of 35-125 cm in depth. Figure 21 shows the gaze error distribution based on the proposed parallax compensation model when the observed targets were 95 cm from the user. The results showed that the proposed parallax compensation model can improve the gaze estimation accuracy by a factor of 10 times. For each meter from the calibration plane, the parallax error adds only a gaze error around ±0.20°, i.e. an improvement of 95.79%. Using simulated data, the proposed parallax compensation model has shown very promisingly to the field of eye tracking.

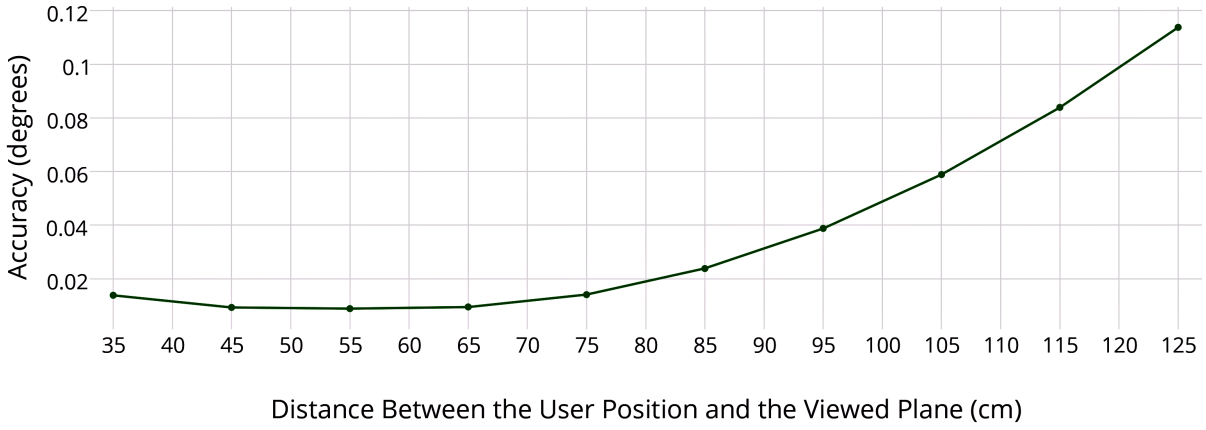


Figure 20: The influence of the proposed parallax compensation model to gaze estimation. The observed targets moved to 10 different distances from the user position, i.e. in a range 35-125 cm in steps of 10 cm. The parallax error only added a gaze error around ±0.02° for each 10 cm from the calibration plane.

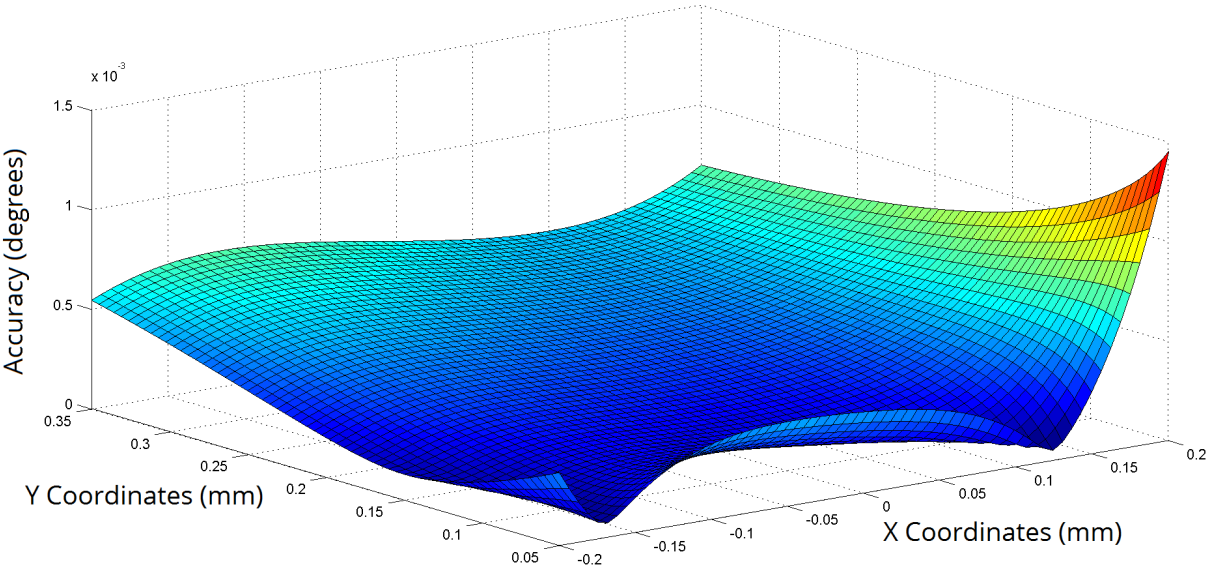


Figure 21: The influence of the proposed parallax compensation model to the gaze estimation. Using the parallax compensation model, the gaze estimation accuracy has achieved 0.04° when the targets have been moved to 95 cm from the user position, i.e. an improvement of 95.79%.

4.4 Assessment using Real Gaze Data

This section presents the evaluation of the proposed parallax compensation model using real gaze data collected in a laboratory experiment. The participants used an uncalibrated head-mounted eye tracker to collect their binocular gaze data. This evaluation aimed to investigate if the proposed model can handle the parallax error in a real eye tracking scenario. The assessment investigated the gaze error distribution concerning (1) magnitude and orientation of gaze error distribution; (2) monocular versus binocular systems; (3) scene camera location; and (4) improvements achieved using the proposed model.

4.4.1 Apparatus

During this assessment, the participants used a similar head-mounted eye tracker prototype like the one built for experiments with Kayak athletes (see Section 7.3). Compared to the eye tracker used for kayak experiments, this one had three scene cameras to investigate the influence of scene camera position in the parallax error. Figure 22 shows a participant using the eye tracker prototype during the parallax error experiment.

Figure 22: The head-mounted eye tracker used for the parallax error experiment. This device had three scene cameras, one eye camera to capture the binocular information from the user (built in May 2017).

This device had the following components (i) four synchronized Full-HD web cameras that allow capturing 30 frames per second with 1280×720 of resolution; (ii) one lens of 3.97mm (focal length) and f/2.8 (aperture) for the eye camera; (iii) three eye-fish lenses of 1.21mm and f/2.0 for the scene cameras; (iv) one IR narrow pass filter; and (v) some off-the-shelf hardware to mount electronic components in a canoe slalom helmet.

4.4.2 Participants

A sample of 20 volunteer participants (15 males and 5 females) was recruited from the IT University of Copenhagen. Four undergraduate students, five master students, seven doctoral students, and four postdocs researchers. Participants ranged from 20 to 45 years old ($mean = 31.6 \pm 7.2$) and their height ranged from 150 cm to 190 cm ($mean = 175.1$ cm $\pm$ 11.2 cm). Eleven had normal vision, six wore glasses, and three wore contact lenses. According to the eye dominance test described by Collins and Blackwell [116], thirteen participants had right-eye dominant and seven left-eye dominant. Participants were free to blink, moving the head, withdraw from testing at any stage.

4.4.3 Procedure

The experiment started with the personal calibration. Figure 23 (*left*) shows the targets used for the personal calibration and the experiment, distributed in a 3×3 matrix on a moving whiteboard. For the personal calibration, the board was 2 meters from the participant. The instructor moved the whiteboard from 2 to 18 meters (see Figure 23 (*right*)) in steps of 2 meters, to collect gaze data in depth. The instructor instructed how long each participant was to fixate on a given target, i.e. around 2 seconds to collect at least 60 gaze samples.

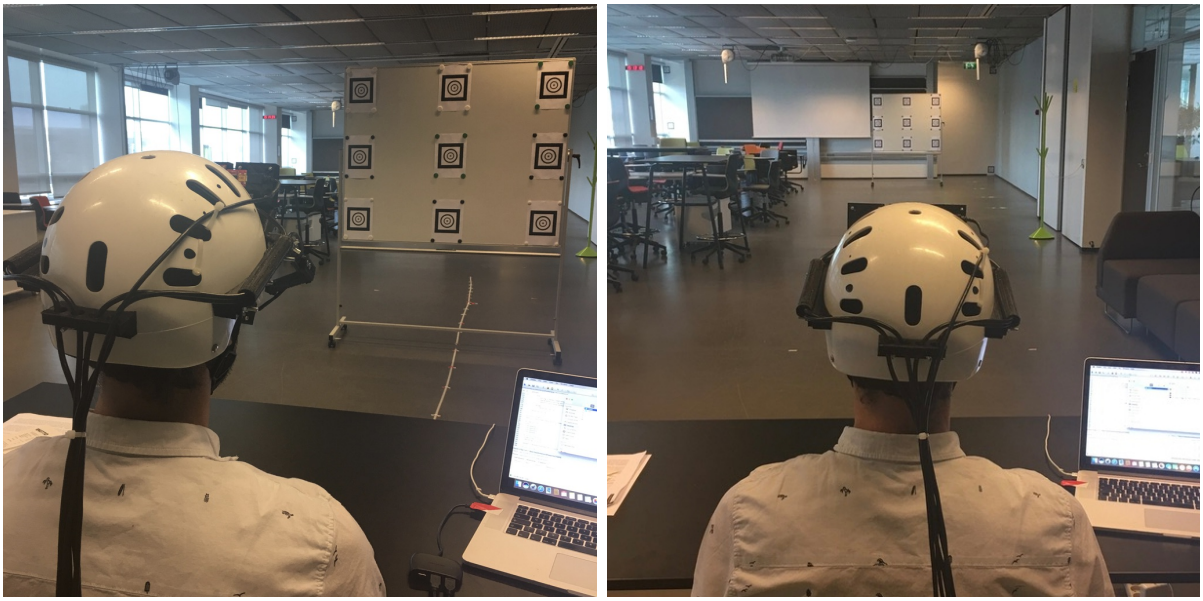


Figure 23: The participants looked at 9 targets on a moving whiteboard to investigate the influence of parallax error and the proposed compensation model when the targets moving in depth. (*left*) the board is 4 meters from the participant; and (*right*) the board is 18 meters from the participant.

4.4.4 Evaluation in a Real Eye Tracking Scenario

First of all, it is important to mention that, as the scene cameras had eye fish lenses, it was not possible to identify the targets automatically because they were small. All targets coordinates (i.e. the red circles presented in all plots of this subsection) were collected manually through mouse clicks in three different scene images. It means the manual data collecting could include outliers in the final results. On the other hand, an eye tracking system provided all gaze estimations automatically (i.e. the colored crosses in the plots).

Magnitude and Orientation of Gaze Error Distribution This experiment investigated if the parallax error distribution would have the same magnitude and orientation among all participants. For each participant, the test evaluated the offset between the gaze estimation and the actual observed target, in 9 depth planes from 2 meters (calibration plane) to 18 meters. This test used only the binocular gaze estimation based on homography [65] from the middle scene camera. Figure 24 shows the gaze error distribution of 4 first participants when gazing 9 targets at 18 meters from them.

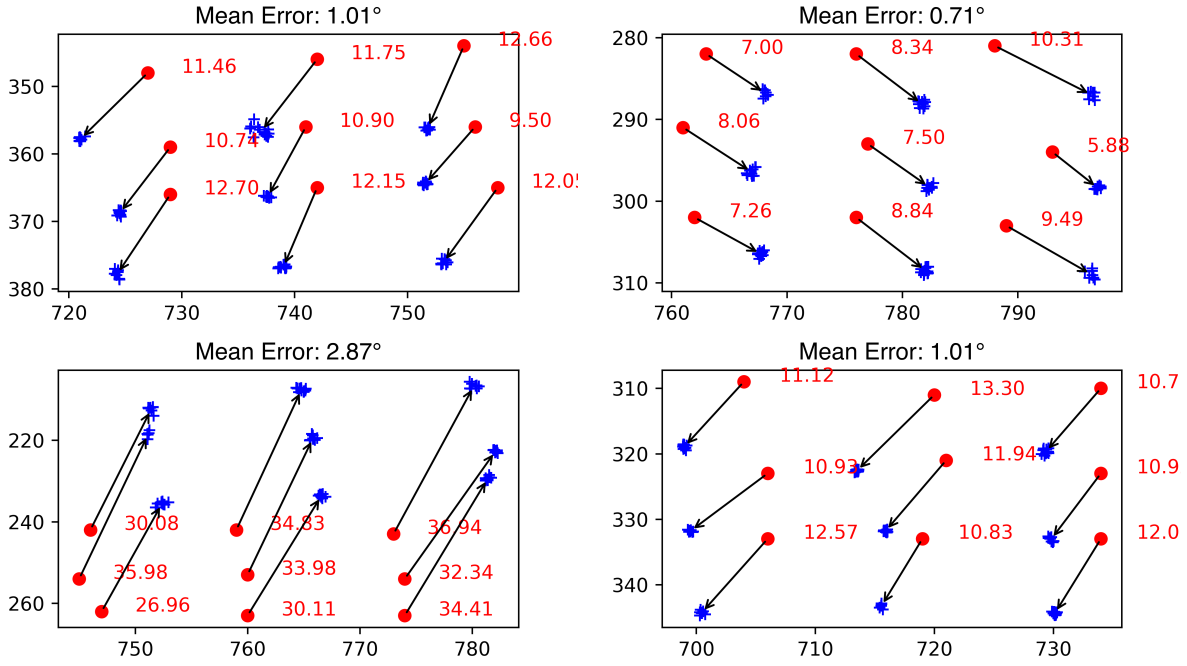


Figure 24: The effect of parallax error in gaze error distribution for 9 observed targets on a depth plane at 18 meters from the participant. Red circles are the targets, blue crosses are the gaze estimations, and red values are the gaze error in pixels. The error distribution of 4 participants are (*top-left*) Participant #01 with left-eye dominant; (*top-right*) Participant #02 with left-eye dominant; (*bottom-left*) Participant #03 with left-eye dominant; and (*bottom-right*) Participant #04 with right-eye dominant.

All participants were at the same position, and the instructor moved the whiteboard in the same direction. However, the gaze error distribution presented different magni-

tudes and orientations among this group of participants. For example, in Figure 24 the gaze estimation of Participant #03 (*bottom-left*) has orientation around 65° and mean magnitude of 2.87°. On the other hand, the gaze estimation of Participant #04 (*bottom-right*) has orientation around 210° (i.e. in the opposite direction) and mean magnitude of 1.02°. It is also possible to notice that the difference between the average magnitude of Participant #03 (*bottom-left*) and Participant #02 (*top-right*) is more than 4 times.

Angle kappa was one of the independent variables did not controlled in this experiment. Angle kappa is different regarding α and β angles between humans [17]. Thus, the difference in magnitude and orientation of parallax error among this group of participants is directly related to subject's angle kappa. For future investigations, it is also important to evaluate the parallax error distribution with respect to other aspects such as (i) the use of uncalibrated head-mounted eye trackers, which the eye camera position is different for each user; (ii) the helmet position related to the participant's head, which captures the eye information in a different angle and gives greater emphasis to one of the eyes; and (iii) the influence of head rotations in the parallax error (see Chapter 5).

It is important to note that the parallax error was almost imperceptible in the processed videos of this experiment. As the head-mounted eye tracker had an eye-fish lens with a wide angle of view (around 220°), the scene camera captures the environment like a sphere and covers the entire subject's field of view. It means, when the subject looks naturally forward, the deeper is the observed plane, the closer it is to the image center, as shown in Figure 25. Of course, the gaze estimation accuracy is better in the calibration plane than in any depth plane. However, 30 pixels of gaze error (in the worst case) is a very small error in high-resolution images. For visual inspector, the gaze estimation will be very close to the actual observed region although the parallax error.

Figure 25: Two different images captured from a scene camera using a M12 eye-fish lens of 1.21mm (focal length) and f/2.0 (aperture). This lens has an angle of view of 220° and it covers the entire subject's field of view. It is also possible to see a small region even behind the subjects. Cameras that use eye-fish lenses with a wide angle of view, capture the environment like a sphere.

Monocular versus Binocular System This experiment evaluated the influence of parallax error in monocular and binocular eye tracking systems. The assessment investigated if there is some difference in using the dominant eye in a monocular setup and the benefits of using binocular gaze estimation. Only a single eye camera captured the binocular data from the participants. The data analysis investigated the left gaze, right gaze, and the average between both eyes as the binocular gaze estimation individually. Figure 26 shows both monocular and binocular gaze error distribution (from the middle scene camera) of 4 first participants when gazing 9 targets at 18 meters from them.

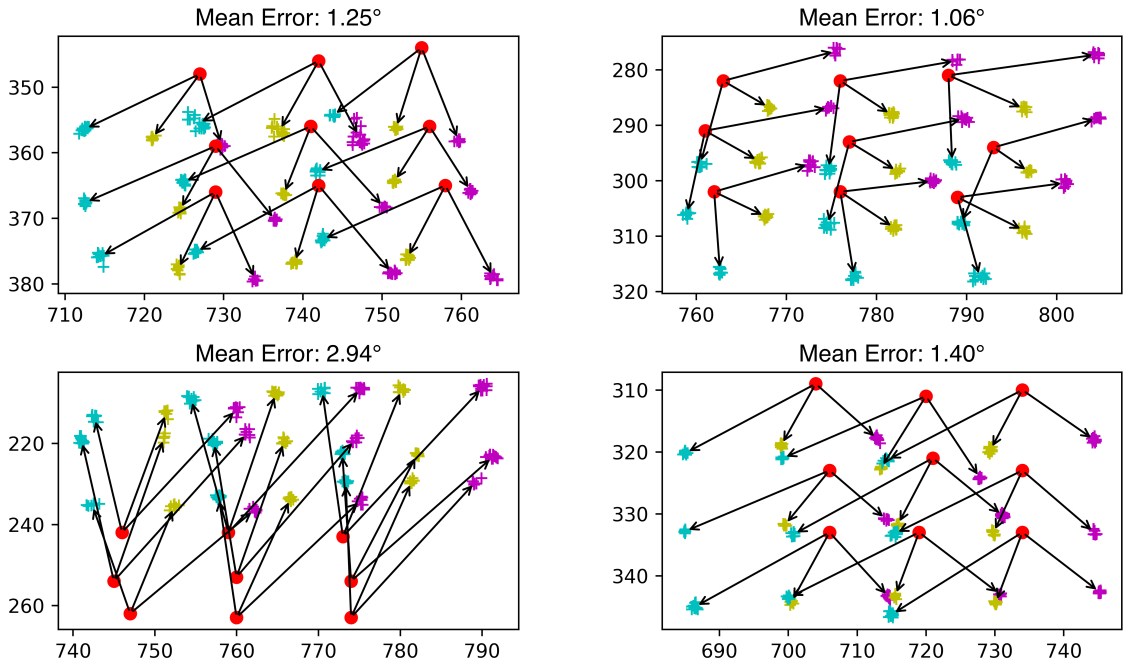


Figure 26: The influence of parallax error in three gaze estimations methods. Magenta crosses are the gaze estimations from the left eye, cyan crosses are the gaze estimations from the right eye, yellow crosses are the binocular gaze estimation (the average of left and right gaze estimations), and red circles are the actual observed targets. All gaze estimation methods are based on homography [65] using 9 targets in the personal calibration. The error distribution of 4 participants are (*top-left*) Participant #01 with left-eye dominant; (*top-right*) Participant #02 with left-eye dominant; (*bottom-left*) Participant #03 with left-eye dominant; and (*bottom-right*) Participant #04 with right-eye dominant.

For all participants and trials, both monocular gaze estimations tended to the same location concerning the binocular gaze estimation. Contrary to the initial hypothesis, the results showed that the monocular gaze estimations from the dominant eye did not provide a more precise estimation (i.e. smallest gaze error in term of magnitude). Both monocular gaze estimations increased the mean absolute error when comparing the results between Figures 25 and 26. In general, the binocular gaze estimation presented the best accuracies. As the middle scene camera and the eye camera were aligned with the participant's face, the average between both eyes provided the best gaze estimation for this device.

Scene Camera Location This experiment evaluated the influence of scene camera location at the head-mounted eye tracker concerning the parallax error. For this reason, the head-mounted eye tracker had three scene cameras in different X-axis locations. All scene cameras used the same type of lenses (eye-fish lens of 1.21mm and f/2.0) to avoid any bias in the collected data. The cameras were in the following positions (i) left-scene camera at 5.1cm to the left; (ii) middle-scene camera at the center of the helmet; and (iii) right-scene camera at 5.1cm to the right. Figure 27 shows the gaze error distributions of the same observed nine targets from three different camera perspectives.

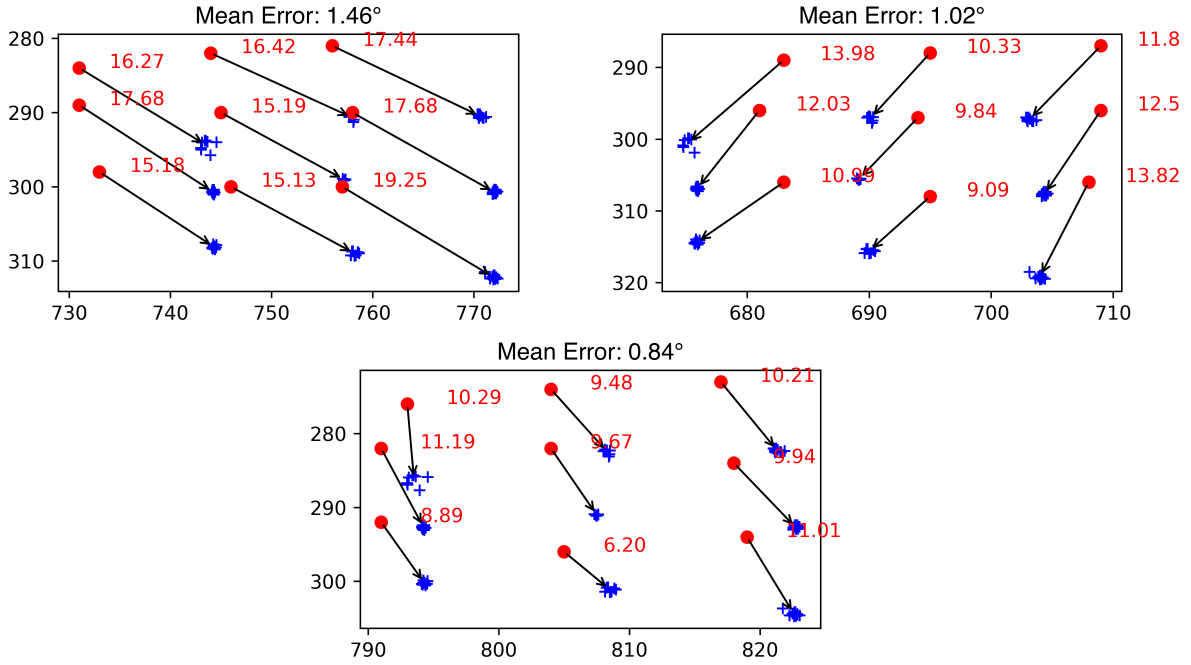


Figure 27: The influence of parallax error in the gaze estimation from three different camera perspectives. Red circles are the actual observed targets, blue crosses are the gaze estimations, and the red values are the gaze error in pixels. The binocular gaze estimations correspond to 9 observed targets at 18 meters from Participant #05. (*top-left*) gaze estimations from left-scene camera; (*top-right*) gaze estimations from right-scene camera; and (*bottom-center*) gaze estimations from middle-scene camera.

The gaze error distribution had a similar behavior for all participants in this experiment. The results showed that the error orientation tends to the left-, middle- and right-scene camera locations. Figure 27 shows the gaze estimation distribution of Participant #05 in a distance of 18 meters from the observed plane. As the scene cameras capture mirrored images from the environment, the error orientation in Figure 27 is in the opposite direction of the X-axis of scene camera. It means the left-most scene camera forces the parallax error orientation to the right, and to the left on the contrary. The results confirmed the hypothesis of Mardanbegi and Hansen [19] that the spatial offset between the user's eyes and the eye tracker cameras influences the parallax error directly.

Improvements achieved using the proposed model During the experiment with real gaze data from 20 participants, the results showed that the parallax error is an user-dependent variable. The magnitude and orientation of gaze error distribution are quite different between this group of participants. Thus, it was not possible to use the proposed parallax compensation model (described in Equation 4.9) in the same way as used during the experiment with simulated data (described in Section 4.3).

As the parallax error changes the *principal point* p in the pinhole model, and it is necessary to use the translation $t_p(p_x, p_y)$ of principal point as priors in the camera matrix (see Equation 4.4). The translation is linear and increases equally for each meter in depth. Euclidean distance d between the average of observed targets and the average of gaze estimations determine the translation parameters in a specific depth Z. To calculate the translation parameters for each meter, it is necessary to divide d by the depth information Z, $p_x = d_x/Z$ and $p_y = d_y/Z$. Figure 28 shows the gaze estimations (from the middle scene camera) of 4 first participants with (*green* crosses) and without (*blue* crosses) using the proposed model, when the participants gazing 9 targets at 18 meters from them.

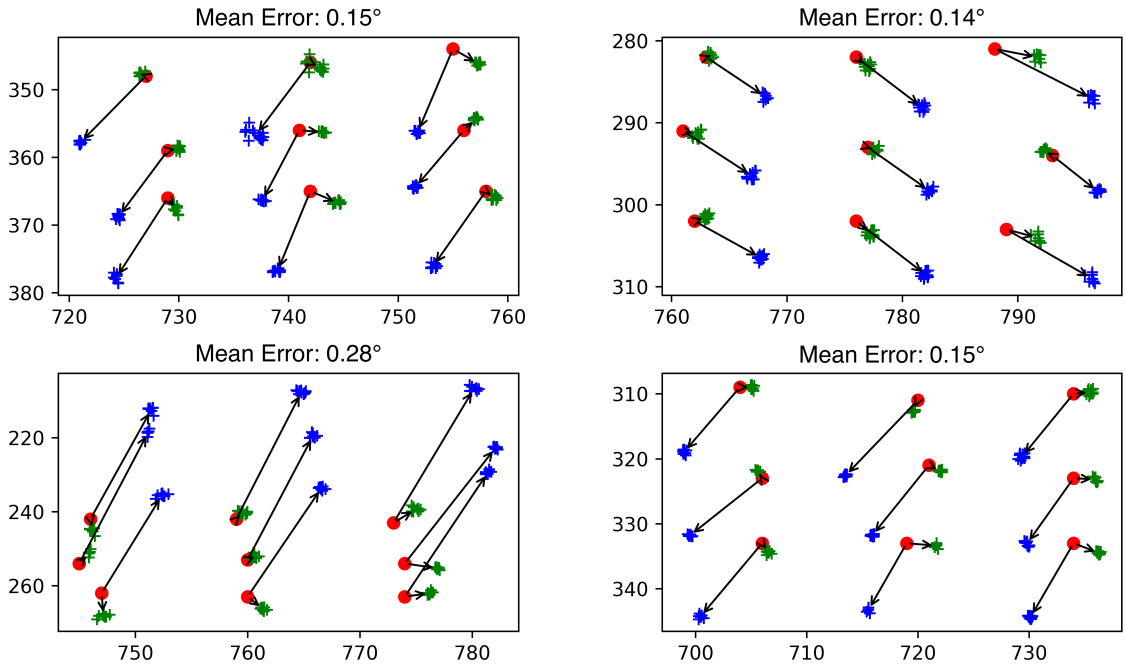


Figure 28: Results achieved using the proposed parallax compensation model. Blue crosses are the binocular gaze estimation based on homography [65], green crosses are the compensated gaze estimations, and red circles are the actual observed targets. The error distribution of 4 participants are (*top-left*) Participant #01 with left-eye dominant; (*top-right*) Participant #02 with left-eye dominant; (*bottom-left*) Participant #03 with left-eye dominant; and (*bottom-right*) Participant #04 with right-eye dominant.

The results in Figure 28 present a huge improvement in the accuracy of gaze estimation in comparison with the results in Figure 24. Figures 29 and 30 show the average gaze

error over all 20 participants of this experiment. The red bars show the gaze estimation under the influence of parallax error, which the average gaze error is $1.01° \pm 0.77°$. The green bars show the results of proposed parallax compensation model using the depth information Z and the translation of principal point t_p as priors, which the average gaze error is $0.20° \pm 0.11°$. The proposed parallax compensation model showed very promisingly to reduce the influence of parallax error in uncalibrated head-mounted eye trackers, and it has presented an improvement of 80.59% in a real eye tracking scenario.

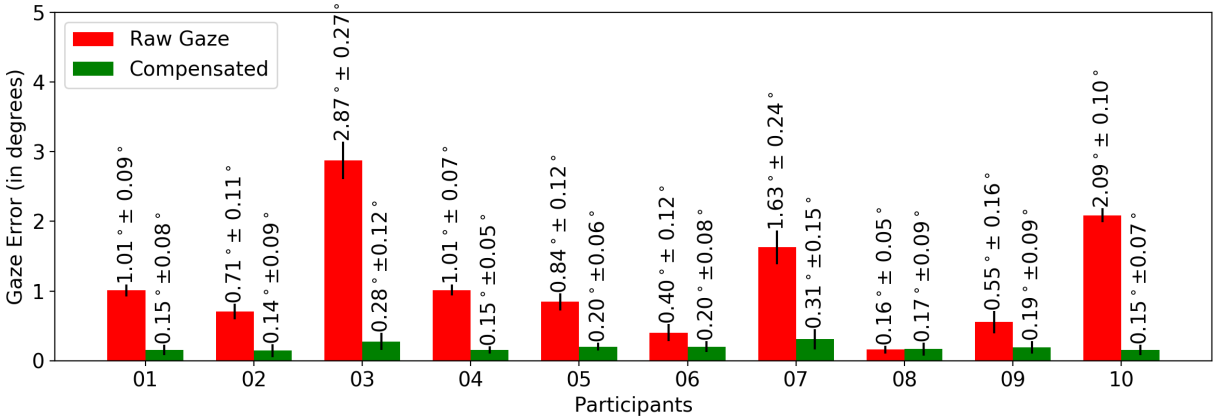


Figure 29: Mean gaze error between all observed targets at 18 meters from the Participants #01 to #10. (*red bars*) Mean gaze estimation using a traditional gaze estimation method based on homography [65]. (*green bars*) Mean gaze error of proposed parallax compensation model.

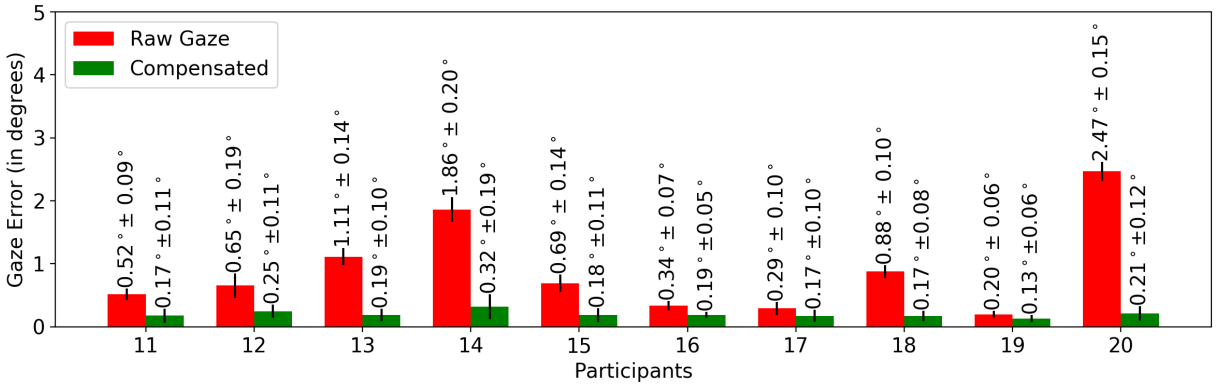


Figure 30: Mean gaze error between all observed targets at 18 meters from the Participants #11 to #20. (*red bars*) Mean gaze estimation using a traditional gaze estimation method based on homography [65]. (*green bars*) Mean gaze error of proposed parallax compensation model.

4.5 Conclusions

This chapter presents a compensation model to minimize the influence of parallax error in head-mounted eye trackers. The proposed model uses two information as priors, namely

(1) the *distance* between the subject and the observed target; and (2) the *offset* of gaze error distribution in any depth plane. The parallax error compensation is based on pure translation motion, which it is a geometric transformation where there is no rotation. Thus, the depth information Z in Equation 4.9 adjusts the gaze data distribution to estimate more accurate subject's gaze even in different depth planes.

The assessment of proposed parallax compensation model used both simulated and real gaze data. The evaluation using simulated gaze data showed that the phenomenon of parallax adds an error around $\pm 2.30°$ for each meter in depth with respect to the calibration plane. On the other hand, using the proposed model the parallax only adds a gaze error around $\pm 0.20°$ for each meter in depth. It means that the proposed model improves the gaze estimation accuracy by a factor of 10 times. The data analysis showed an improvement of 95.79% in a simulated scenario.

The assessment also used real gaze data to investigate if the proposed model can handle the parallax error in a real eye tracking scenario. The data analysis of 20 participants showed that the magnitude and orientation of gaze error distribution are different for each subject due to the angle kappa is an user-dependent variable. For simulated experiment, it was necessary only use the depth information Z as priors to compensate the parallax error. However, as the subject's angle kappa changes the *principal point* p in the pinhole model, the proposed compensation model needs the angle kappa offset to handle the parallax error in a real eye tracking scenario.

The results also showed that it is possible to compensate the influence of parallax error for both monocular and binocular eye tracking systems. For example, in Figure 26, the monocular (*magenta* and *cyan* crosses) and binocular (*yellow* crosses) gaze error have basically the same distribution but in a different location (i.e. including a translation). Thus, it is necessary to investigate the gaze error distribution in depth for each subject. During the personal calibration, the subject must also observe a set of targets on a depth plane to verify the influence of angle kappa in the gaze error distribution in depth. The minimum number of observed targets necessaries to identify the influence of angle kappa in depth planes will be a research topic in a future investigation.

The evaluation using real gaze data showed that the phenomenon of parallax adds a gaze error around $1.01° \pm 0.77°$ when the targets move 16 meters in depth concerning the calibration plane. On the other hand, using the proposed compensation model the parallax only adds a gaze error around $0.20° \pm 0.11°$ for the same depth plane. The data analysis showed an improvement of 80.59% in a real eye tracking scenario.

This chapter also describes a gaze estimation method based on epipolar geometry. The method uses binocular information to define two epipolar lines in the scene image, and the gaze estimation is based on the intersection between the left and right epipolar lines. The proposed gaze estimation method achieved high accuracy estimation in simulated data ($error < 0.01°$). However, this gaze estimation method is very sensitive to feature detection noise and its accuracy is similar to others traditional gaze estimation methods in a real eye tracking scenario. As fundamental matrix based on the eight point algorithm [111] is sensitive to noise [115], the proposed gaze estimation method did not present good performance in real head-mounted eye tracking scenarios, specially due to complex eye movements (e.g. torsions) and natural subject's head rotations. For this reason, all experiments in real scenarios presented in this Ph.D. thesis used a homography-based eye tracking method with 9 calibration targets [65].

The assessment of proposed parallax compensation model using simulated gaze data was accepted as a full paper[1] in 2^{nd} IEEE International Workshop on Computer Vision in Sports (CVsports 2015) of 2015 IEEE International Conference on Computer Vision (ICCV 2015) and is available in Annex A, and as an abstract in 19^{th} European Conference on Eye Movements (ECEM 2017) and is available in Annex B. The assessment using real gaze data will be submitted as a full paper to 2018 Symposium on Eye Tracking Research & Applications (ETRA 2018), and a more detailed data analysis (e.g. using gaze data from all depth planes, investigating the minimum number of observed targets on a depth plane) will be submitted as a full paper to some journal in the field of computer science.

[1] http://dx.doi.org/10.1109/ICCVW.2015.107

Chapter 5

Using Priors to Compensate the Head Rotation Error

This chapter focuses on compensating the influence of head rotations in the gaze estimation. In general, head-mounted eye trackers are negatively influenced by head rotations because these movements add a significant spatial error on the gaze estimation, in the opposite direction of the head rotation. The eye tracking system needs to be robust to this kind of head movements; otherwise, the analysis of gaze data will be impracticable.

This chapter proposes a head rotation compensation model that uses the three-dimensional angles from the subject's head as priors to decrease the rotational errors. It is possible to obtain the updated head angles using an orientation sensor attached to the head-mounted eye tracker. The proposed compensation model is based on least squares fitting between the gaze estimation and current head pose with the aim to maintain high accuracy and precision of the eye tracking system along the entire session.

Section 5.1 describes the spatial error produced by the head rotations in more detail, and the proposed head rotation compensation model is described in Section 5.2. The evaluation of the proposed model using real eye tracking data from 10 participants in a laboratory experiment is described in Section 5.3. Section 5.4 explains the entire personal calibration used for the gaze estimation method and for both head rotation and parallax error compensation models. Finally, in Section 5.5 are presented the final conclusions of the proposed head rotations compensation model.

5.1 Influence of Head Rotations in Gaze Estimation

Head-mounted eye trackers are naturally robust to user's head movements, because all their components (i.e. eye camera, scene camera, infrared light sources, frame, helmet,

sensors) move together following the user's head movements [17]. Figure 31 shows the gaze estimation as a blue circle that covers the current observed target on the calibration plane. After the personal calibration, it is possible to estimate the user's gaze with high accuracy (i.e. less than 0.50°) on the calibration plane.

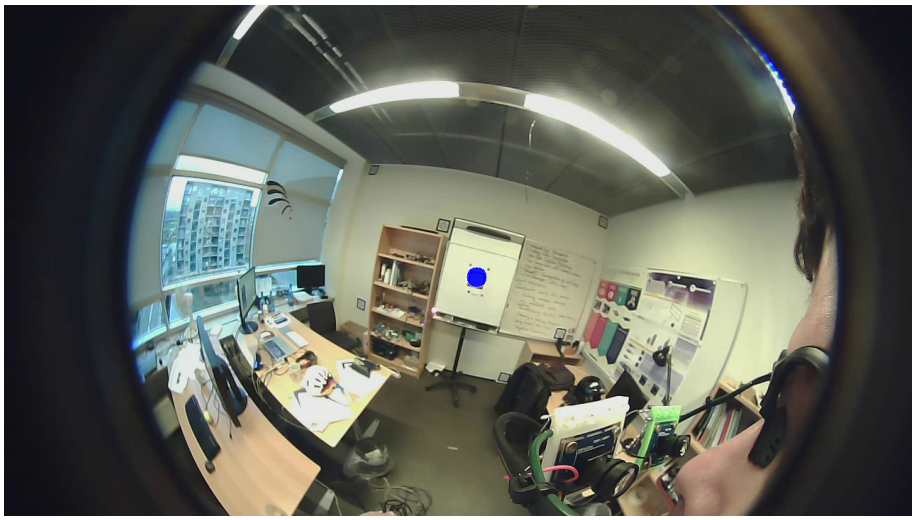


Figure 31: High accuracy gaze estimation of head-mounted eye tracking on the calibration plane using a binocular gaze estimation method based on homography and seven calibration targets [65].

During some experiments, we observed the influence of natural head rotation in the gaze estimation, especially movements around *roll* axis (i.e. Z-axis in the right-hand rule). Figures 32 (*left*) and (*right*) show the spatial error when the subject keeps his focus of attention in a calibration target and rotates the head around Z-axis. It is important to note that this spatial error does not suffer any influence of parallax error, once the observed target is still at the same distance on the calibration plane.

Figure 32: The spatial error produced by the user's head rotations to the left and the right around Z-axis.

As head movements do not change the rotation matrix properties of pinhole model (see Equation 4.6) in head-mounted eye trackers, the investigation of the cause of the rotational

errors analyzed the eyes movements behavior during head rotations. An observational analysis identified that eyes perform complex movements when the subject keeps the gaze in a fixed target and rotates the head. These movements are similar to ocular torsions, which the pupil center coordinates change a lot compared to the original head position. Figure 33 shows the binocular pupil center coordinates from two subjects when performing head rotations to the left (*cyan crosses*) and the right (*magenta crosses*).

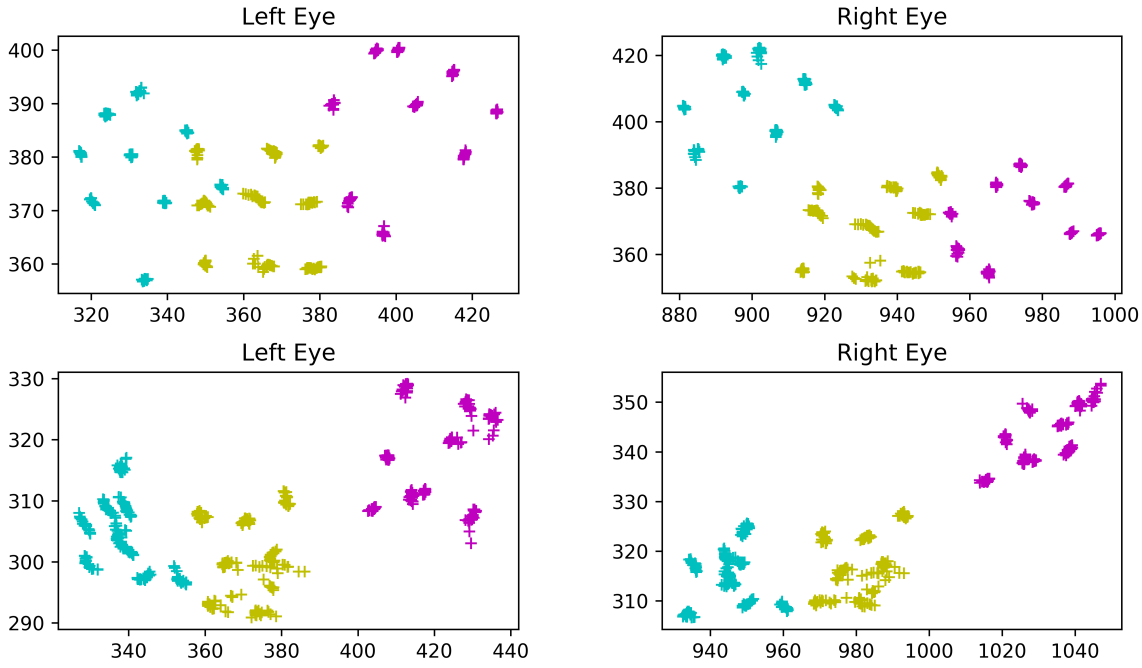


Figure 33: The effect of head rotation in both pupil center coordinates from two subjects. Cyan crosses are the pupil center coordinates when the subject rotates the head to the left. Yellow crosses are the pupil center coordinates in the original head position; Magenta crosses are the pupil center coordinates when the subject rotates the head to the right.

During this task, two subjects looked at nine targets on a plane and rotated their head to the left and the right for each observed target. Figure 33 shows a similar pupil center distribution as a 3×3 matrix for both eyes. It is also possible to observe that the head rotation creates three different classes of pupil centers. Because of the changes in the pupil center coordinates in the eye image, the gaze estimation presents a significant offset concerning the actual observed target in the scene image.

It is easier to observe this phenomenon directly in the eye images. Figure 34 (*left*) shows both eyes when the participant rotates his head to the left, and Figure 34 (*right*) shows an ocular torsion when the participant rotates his head to the right. Note the difference in the pupil center in both eye images, although the observed target is in a very similar position in the scene images. As head rotation changes the eye feature coordinates,

eye tracking methods would wrongly estimate the subject's gaze.

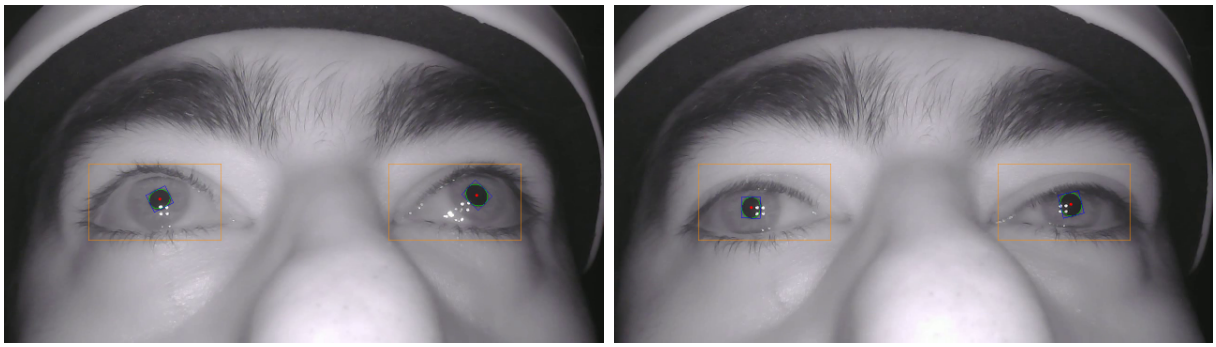


Figure 34: During the head rotation around Z-axis, the eyes perform a more complex torsion movement (e.g. convergence, divergence). (*left*) shows the eyes position when participant looks at a fixed target on the calibration board and rotates his head to the left. On the other hand, (*right*) shows different pupil center positions for the same observed target, but with head rotation to the right.

5.2 Head Rotation Compensation Model

Head rotations are natural human movements controlled by the activities of muscles in the region of the neck. According to Alfayad et al. [117], the human neck motions and head rotations are classified based on three major movements, namely (i) *extension* and *flexion*, pitch rotations around X-axis; (ii) *vertical rotation*, yaw rotations around Y-axis; and (iii) *lateral bending*, roll rotations around Z-axis. Figure 35 shows in details the human head rotations relative to Euler angles, 3D axes and maximum angle of motion.

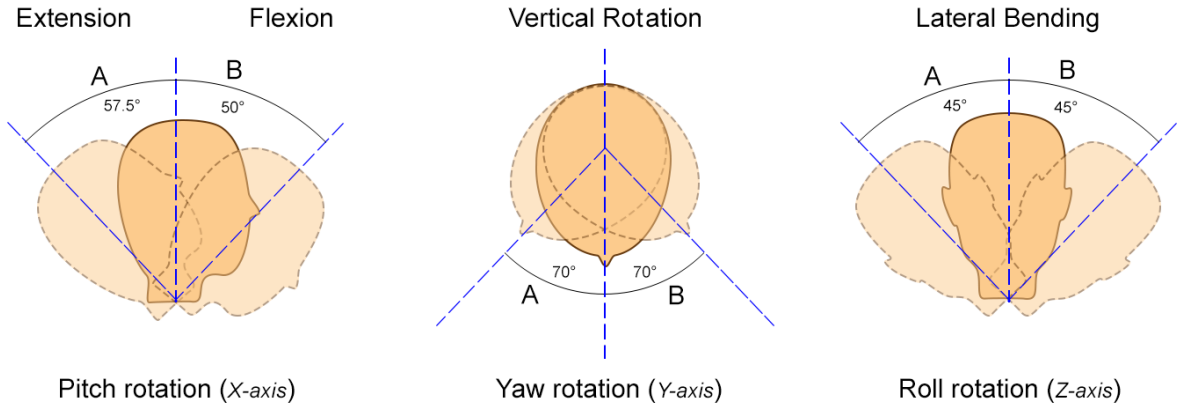


Figure 35: Human head rotations based on three major neck motions. (*left*) shows the extension and flexion motion around pitch axis in Euler angles. (*center*) shows the vertical rotation motion around yaw axis in Euler angles. (*right*) shows the lateral bending motion around roll axis in Euler angles.

Flexion motion allows the head to bend toward the chest in a range of 50°, while *extension* motion ensures the head to tilt back to the primary position in a range of 57.5° [117]. These motions occur when a subject shakes the head to say *yes*. *Vertical rotation* motion allows the head to turn to the right and the left in a range of 70° [117], and occur when a subject shakes the head to say *no*. In turn, *lateral bending* motion allows the head to bend towards the shoulders in a range of 45° [117]. This motion is commonly used to stretch the neck, and it has the biggest influence to the gaze estimation.

The proposed compensation model uses the head angles as priors to compensate the influence of head rotation around (X, Y, Z)-axes in the gaze estimation. In this case, an orientation sensor attached to the structure of head-mounted eye tracker provides the information about head movements. The sensor was installed behind the scene camera, and it provides the updated rotation matrix and the three-dimensional subjects' head movements in Euler angles (yaw, pitch and roll). As the error distribution is user-dependent, it is necessary to perform a personal calibration exclusively for head rotation compensation.

During the personal calibration process, the subjects must look at a set of targets, and rotate their head to the left and the right. Then the magnitude and orientation $\alpha(x_h, y_h)$ of the difference between the gaze estimation and observed target are computed. The head rotation compensation model uses a *least squares polynomial* to fit the gaze error α related to the angles from one of the head rotation axes β, according to Equation 5.1:

$$\alpha^{'} = \sum_{i=0}^{k} \beta_i \alpha_i. \tag{5.1}$$

Finally, the gaze estimation without the influence of head rotation can be estimated through the subtraction of the predicted gaze error $\alpha^{'}$ (based on the current head rotation angle β_i) and the gaze estimation g, as described in Equation 5.2:

$$g^{'} = g - \alpha^{'}. \tag{5.2}$$

5.3 Assessment using Real Gaze Data

The evaluation of the proposed head rotation compensation model used gaze data from a real eye tracking scenario in a laboratory experiment. The participants used an uncalibrated head-mounted eye tracker to collect binocular information. Therefore, the evaluation investigated if the proposed compensation model can handle the influence of head rotation in an eye tracking application. The evaluation investigated the error distribution

in gaze estimation concerning (1) magnitude and orientation of gaze error distribution; (2) monocular versus binocular systems; (3) scene camera location; and (4) improvements achieved using the proposed model.

5.3.1 Apparatus

The participants used the same head-mounted eye tracker prototype built for the parallax error experiment (see Subsection 4.4.1). The only difference is the use of an orientation sensor to measure head movements. The eye tracker had the following components (i) four synchronized Full-HD web cameras that allow capturing 30 frames per second with 1280×720 of resolution; (ii) one lens of 3.97mm (focal length) and f/2.8 (aperture) for the eye camera; (iii) three eye-fish lenses of 1.21mm and f/2.0 for the scene cameras; (iv) one IR narrow pass filter; (v) one orientation sensor to monitor the head rotations; and (vi) some off-the-shelf hardware to mount electronic components in a canoe slalom helmet.

5.3.2 Participants

A sample of 10 volunteer participants (8 males and 2 females) was recruited from the IT University of Copenhagen. Three undergraduate students, one master students, four doctoral students, and two postdocs researchers. Participants ranged from 20 to 43 years old ($mean = 30.5 \pm 7.6$) and their height ranged from 165cm to 188cm ($mean = 175.4$cm $\pm$ 7.7cm). Six had normal vision, two wore glasses and two wore contact lenses. According to the eye dominance test described by Collins and Blackwell [116], six participants had right-eye dominant and four left-eye dominant. Participants were free to blink, move the head, withdraw from testing at any stage.

5.3.3 Procedure

The experiment started with the personal calibration, using nine targets on a whiteboard located at four meters from the participant. The participants kept their heads still during the calibration data collection, and they stood in front of the calibration board during the entire eye tracking session. The instructor asked the participants to look at each calibration target again, and turn their heads (i) to the left, (ii) then to the primary position, and (iii) finally to the right. The instructor controlled how long the participant looked at each target by head rotation, to collect at least 60 gaze samples (i.e. around 2 seconds). In the end, the participants performed a recalibration to compare the gaze

estimation at the beginning and the end of the eye tracking session. Figure 36 shows a participant during the head rotation experiment.

Figure 36: The participants were asked to look at nine targets during the head rotation experiment. (*left*) the participant turns his head to the left; and (*right*) the participant turns his head to the right.

5.3.4 Evaluation in a Real Eye Tracking Scenario

Magnitude and Orientation of Gaze Error Distribution This experiment investigated if the error distribution produced by head rotation would have the same magnitude and orientation among all participants. For each participant, the experiment evaluated the gaze error in 9 different observed targets at 4 meters away from the participants. Only the binocular gaze estimations from the middle scene camera was assessed to avoid bias relative to the camera location. Figure 37 shows the gaze error distribution of 4 first participants (*each row*) when they turned their heads in lateral bending to the left (*left column*) and to the right (*right column*).

Although all participants stood in the same position at 4 meters from the targets, the gaze error distribution presented a different behavior for each participant. For example, Figure 37 shows a big variance between error magnitude of Participant #01 [*first row*] (i.e. 5.72°) and Participant #03 [*third row*] (i.e. 14.14°) when they turned their heads to the left [*left column*]. The orientation also presented different behavior between the participants. However, it is possible to note the error distribution tends to the opposite direction concerning the lateral bending.

This experiment showed that the rotational error is user-dependent, and the eye tracking system needs to investigate the error behavior for a specific subject. It is important to note the need to perform a personal calibration for the head rotation with the aim to create the regression mapping used to compensate the influence of head rotations.

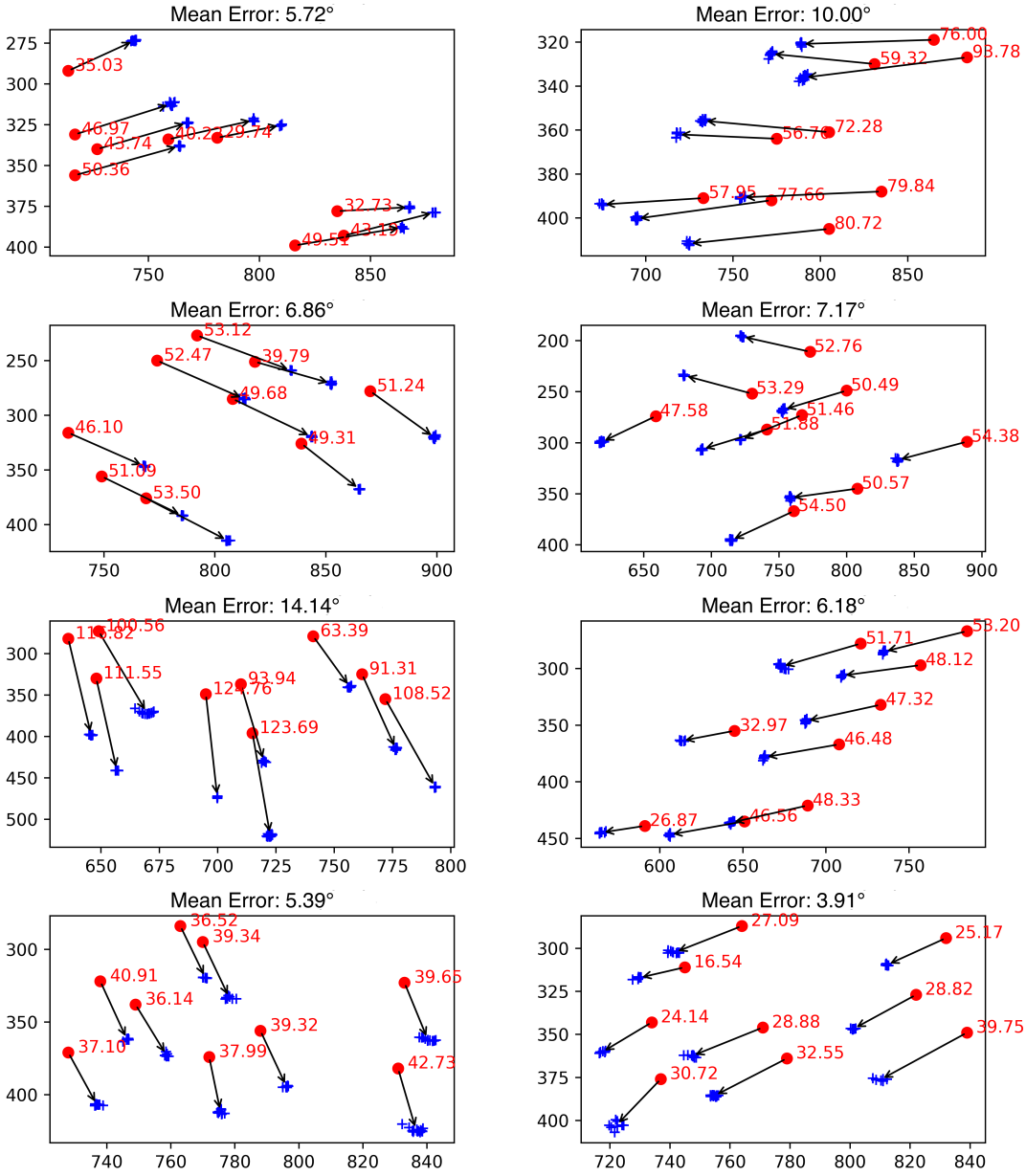


Figure 37: The effect of head rotation in gaze error distribution for 9 observed targets at four meters. Red circles are the actual observed targets, blue crosses are the gaze estimation, and red values are the gaze error in pixels. Each row corresponds to one participant. (*Left collumn*) shows the error distribution when the participants turn their heads to the left and (*right collumn*) shows the roll rotations to the right.

Monocular versus Binocular System This experiment evaluated the influence of head rotations in monocular and binocular eye tracking systems. The assessment investigated if the proposed compensation model can work in both eye tracking systems. The data

analysis assessed separately the left, right and binocular (average between both eyes) gaze estimation. Figure 38 shows the gaze error between the actual observed target and the three gaze estimations from 3 first participants.

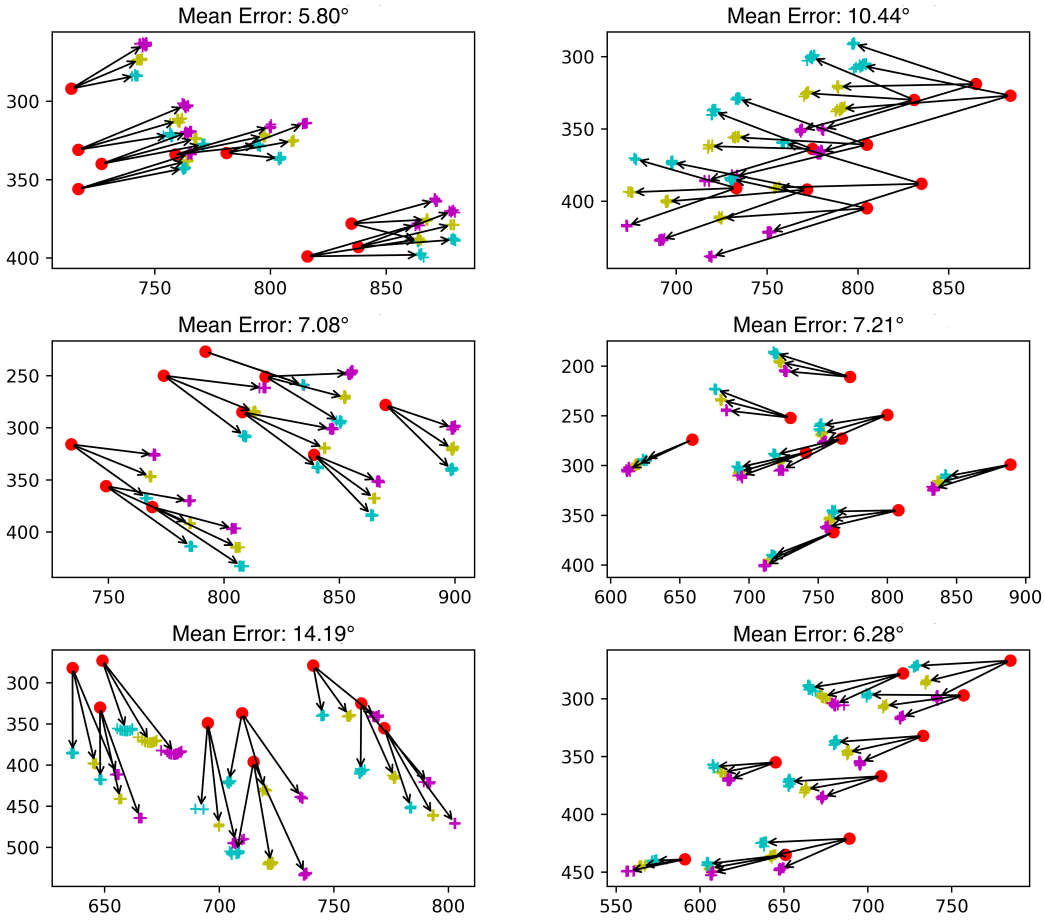


Figure 38: Error distribution of head rotations for 3 gaze estimation methods. Magenta crosses are the gaze estimations from the left eye, cyan crosses are the gaze estimations from the right eye, yellow crosses are the binocular gaze estimation (average of left and right gaze estimation), and red circles are the observed targets. Each row corresponds to one participant. (*Left collumn*) shows the error distribution when the participants turn their head to the left and (*right collumn*) shows the roll rotations to the right.

For all participants, the results showed a strong relationship between the three gaze estimations regarding magnitude and orientation. The gaze error tends to the opposite direction concerning the head rotation. It is important to note an inversion phenomenon in the left and right gaze estimations when the subject turns the head in opposite lateral bedding. Figure 38 shows the left gaze estimation (magenta crosses) is above the others estimations in the left column, and under in the right column (the same phenomenon occurs with cyan crosses). This behavior occurs independently of the dominant eye.

The results also showed that the magnitude and orientation of each gaze estimation method presented similar behavior compared the Euclidean distance between the esti-

mated gazes (colored crosses) and the observed targets (red circles). This experiment showed that it is possible to compensate the influence of head rotations for both monocular and binocular eye tracking systems. It is only necessary to investigate the gaze error behavior through a specific personal calibration.

Scene camera location This experiment evaluated the influence of scene camera position in the eye tracker structure concerning the error produced by head rotations. The head-mounted eye tracker had three scene cameras at a distance of $5.1cm$ in x-axis between their optical center. Figure 39 shows the error distribution (magnitude and orientation) of one participant when she has turned her head to the left in roll rotation.

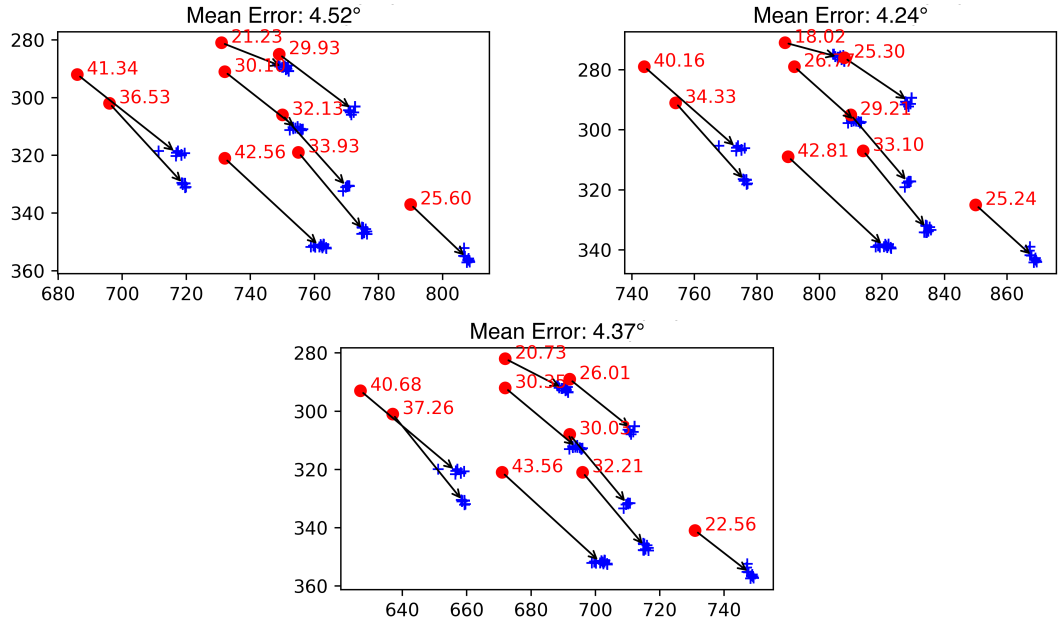


Figure 39: The influence of head rotations in three different scene images. Red circles are the actual observed targets, blue crosses are the gaze estimations, and red values are the gaze error in pixels. (*top-left*) gaze estimations from left-scene camera; (*top-right*) gaze estimations from right-scene camera; and (*bottom-center*) gaze estimations from middle-scene camera.

Unlike the parallax error experiment, which the scene camera position determined different gaze error behaviors (see Subsection 4.4.4), the experiments involving head rotations showed there is no influence of camera location in the orientation of gaze error. For all participants, the orientations and magnitude are very similar to each other, even in different scene images. The small difference in the error magnitude is because the targets coordinates were collected manually through mouse clicks in three different scene images.

Improvements achieved using the proposed model During the experiment with real gaze data from 10 participants, the results showed that the influence of head rotation is an user-dependent variable. The lateral bending motion around Z-axis adds an average

error of $1.92° \pm 0.79°$ in the gaze estimation for each 10° of head rotation. The average gaze error magnitude was $9.01° \pm 5.18°$ when the participants turned their heads to the left, and $7.83° \pm 3.67°$ to the right. The proposed compensation model reduced the gaze error respectively to $1.25° \pm 0.73°$ (*left*), and $1.06° \pm 0.50°$ (*right*). Figure 40 shows the gaze estimations with (green crosses) and without (circle crosses) using the proposed head rotation compensation model, when the first 4 participants turned their head to the left (*left column*) and the right (*right column*).

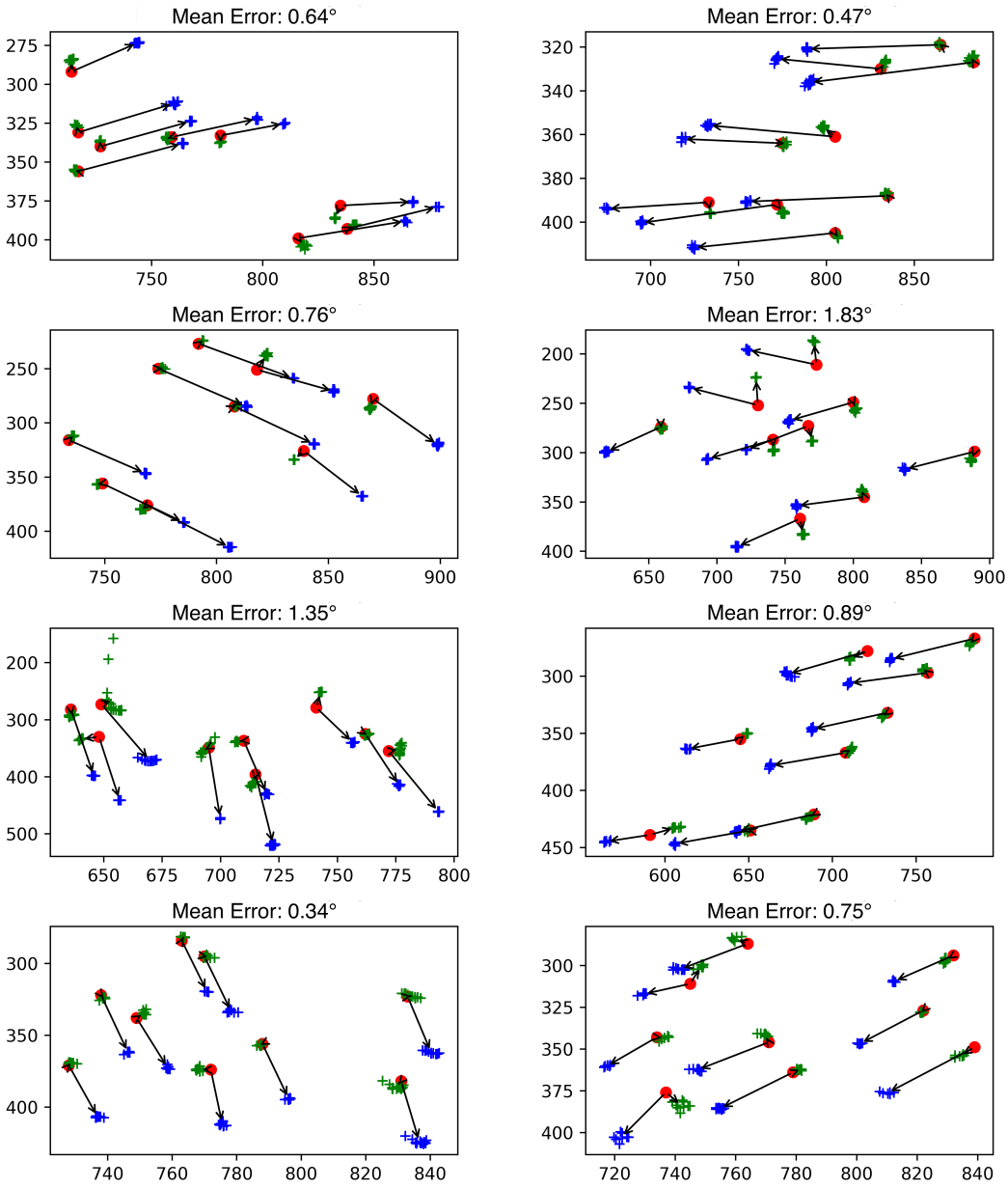


Figure 40: Results achieved using the head rotation compensation model. Blue crosses are the binocular gaze estimation based on homography [65], green crosses are the compensated gaze estimations, and red circles are the actual observed targets. (*left collumn*) shows the gaze error distribution when the participants turned their head to the left and (*right collumn*) shows the head rotations to the right.

Figures 41 and 42 show the average gaze error over all 10 participants of this experiment, when they turned their head respectively to the left and the right. The red bars present the gaze estimation under the influence of head rotation, and the green bars present the results of proposed head rotation compensation model using the head angles as priors. The head rotation angles ranged from 16.15° to 68.99° (average 44.47° ±12.96°), and this is the cause of the significant variance between the average gaze error magnitude from the participants in this experiment. The proposed head rotation compensation model showed very promisingly to reduce the influence of head rotations in head-mounted eye trackers, and it has presented an enhancement of 86.41% in a real eye tracking scenario.

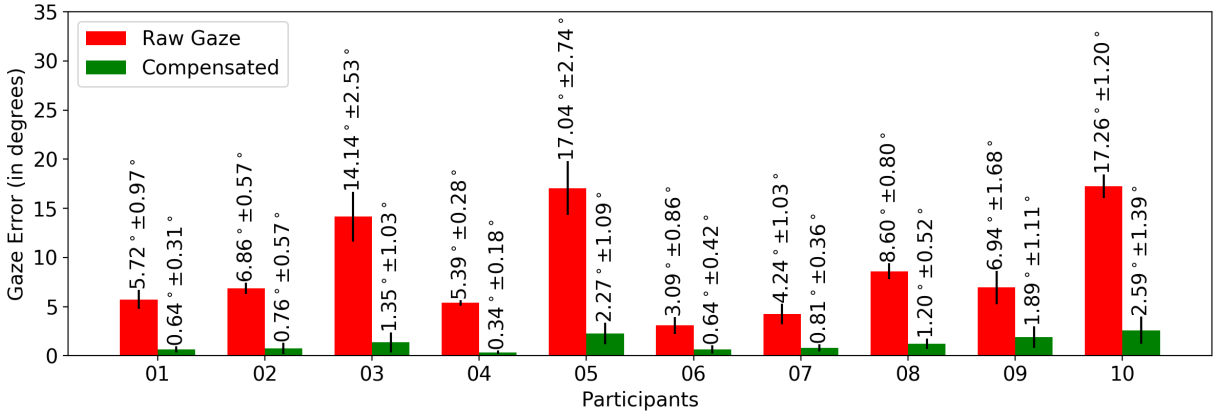


Figure 41: Average gaze error of all observed targets when the participants turned their heads to the left. (*red bars*) Average gaze estimation using a traditional gaze estimation method based on homography [65]. (*green bars*) Average gaze error of proposed head rotation compensation model.

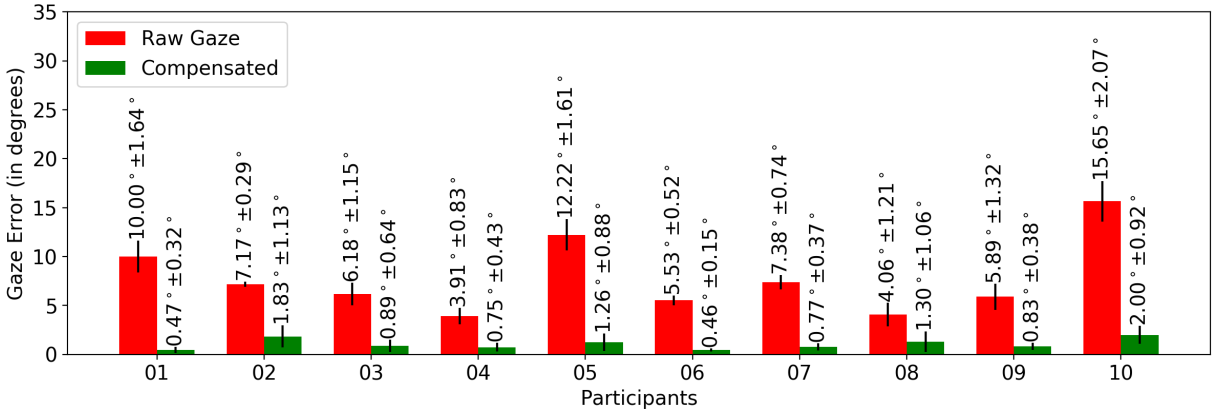


Figure 42: Average gaze error of all observed targets when the participants turned their heads to the right. (*red bars*) Average gaze estimation using a traditional gaze estimation method based on homography [65]. (*green bars*) Average gaze error of proposed head rotation compensation model.

Figure 43 shows the use of proposed model to compensate the influence of head rotation in real gaze data. It shows the results achieved (green circles) from different lateral bending movements when the participant has looked at nine targets in a whiteboard.

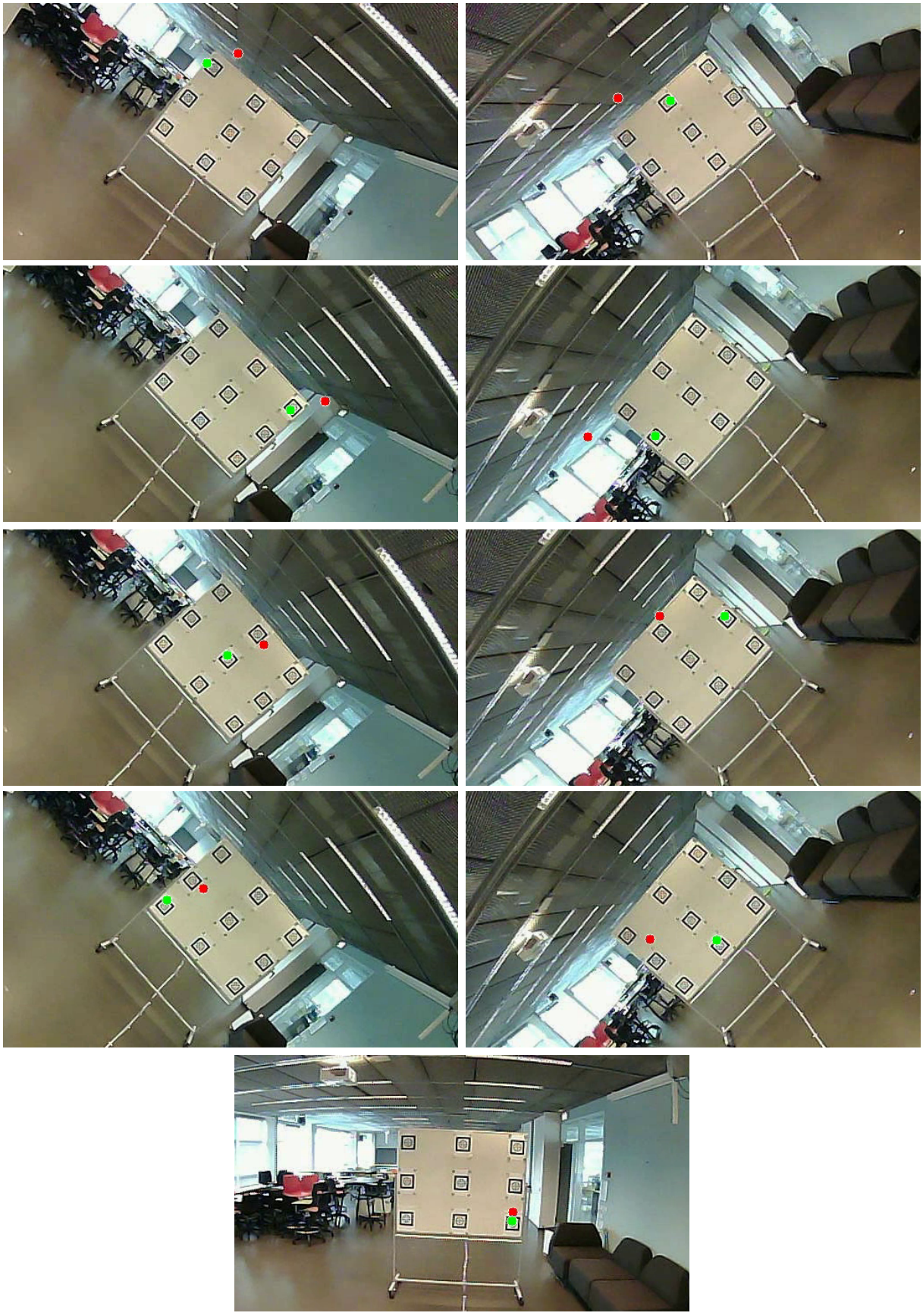


Figure 43: Results achieved using the head rotation compensation model. The red circles illustrate the raw gaze estimations. On the other hand, the compensated gaze estimations are the green circles. Each figure shows one sample of nine targets observed by the participant during this experiment.

5.4 Proposed Personal Calibration

Both parallax error and head rotation compensation models have user-dependent variables. Thus, it is necessary to investigate the gaze error compensation for each subject with the aim of training a set of methods used to compensate both geometrical problem.

The proposed personal calibration is divided into three steps, namely (1) perform the personal calibration in the traditional way, which the user looks at a set of targets on the calibration plane; (2) rotate the head looking at a fixed target to learn person specific parameters; (3) move the calibration target in depth, perform a new calibration with the aim to estimate the user's angle kappa.

To enjoy the benefits proposed in this Ph.D. thesis, after performing the three calibration steps, the eye tracking system must execute following actions (i) gaze estimation, using pupil center coordinates as input data; (ii) head rotation compensation, using head angles from an orientation sensor as priors; and (iii) parallax error compensation, using the distance between the user and the observed target as priors.

5.5 Conclusions

This chapter presents a compensation model to minimize the influence of head rotations in head-mounted eye trackers. The proposed model uses the three-dimensional angles from the subject's head as priors. An orientation sensor attached on the head-mounted eye tracking (behind the scene camera) provides the updated rotation matrix and the subject's head angles. It is important to note the error orientation tends to the opposite direction concerning the subject's head rotation.

The proposed head rotation compensation model is based on a second-order ordinary least squares polynomial. Thus the head angle β in Equation 5.1 predicts the magnitude and orientation of gaze error for a specific head pose, then adjusts the actual gaze estimation in Equation 5.2. As the gaze error distribution is user-dependent, it is necessary to learn the error behavior during the personal calibration.

Unfortunately, it was not possible to assess the proposed model using simulated data, because the simulator does not simulate the binocular eye torsion movements properly. However, the data analysis from 10 participants showed that the head rotation compensation model can be used for both monocular and binocular eye tracking systems, once the analyzed gaze estimation methods (left, right and average binocular) showed a strong

correlation between their error distributions. When the participants turned their heads around Z-axis, the gaze error presented a systematic inversion in left and right gaze estimations. The results did not present any influence of scene camera positions, glasses, contact lenses and dominant eye in the gaze error behavior.

In this dataset, head rotations added an average error around $8.42° \pm 4.73°$ in the gaze estimation. Lateral bending motion added an average error of $1.92° \pm 0.79°$ for each $10°$ of head rotation. On the other hand, the proposed compensation model reduced the average error to $1.15° \pm 0.99°$ and presented an improvement in the accuracy in a head-mounted eye tracking system of 86.41%.

The head rotation compensation model is going to submit as a full paper to 2018 Symposium on Eye Tracking Research & Applications (ETRA 2018).

Chapter 6

Using Priors in Eye Tracking Methods

The use of eye tracking in sports is challenging due to many aspects, such as (i) eye feature detectors non-robust to critical situations; (ii) noises from the environment; (iii) uncontrolled light conditions; (iv) geometric constraints of uncalibrated eye trackers, among others. However, eye tracking is still the best mechanism to supply detailed information about the visual interest of an athlete and to assist sports specialist in the eye movements analysis that cannot be obtained otherwise. Eye tracking systems need to be accurate and robust in several sports situations and, at the same time, to support athletes seamlessly in their daily training sessions and official matches.

Noises from outdoor environments (e.g. light reflections, shadows, sun light) influence negatively head-mounted eye trackers and make impractical the eye feature detection in most sports disciplines in outdoor fields. For this reason, this chapter presents a set of methods that use the geometric relationship between both eyes as priors to improve eye tracking in sports. All eye tracking methods proposed in this chapter embody three main aspects, namely (i) a robust binocular pupil detector; (ii) a novel glint normalization approach; and (iii) an alternative eye camera slippage compensation.

Most of the head-mounted eye tracking systems use the pupil center purely to estimate the subject's gaze. Thus, it is essential to detect precisely the actual pupil center in most processed eye images and, at the same time, avoid changes in the eye camera location during the entire eye tracking session. All methods proposed in this chapter were used to collect and analyze gaze data from experiments in elite sports practices.

Section 6.1 describes a binocular eye feature detector that uses information from both eyes to improve the pupil detection and to make the eye tracking system more robust to changes in the illumination conditions. Two methods to compensate the cameras slippage

concerning the user's faces are described in Section 6.2 (using glints normalization) and Section 6.3 (using eye landmark detection). Finally, Section 6.4 presents the conclusions of proposed eye tracking methods using eye information as priors.

6.1 Binocular Eye Feature Detector

According to Kaufman and Albert [118], six extraocular muscles control the human eye movements, namely: (i) superior rectus muscle (*srm*); (ii) inferior rectus muscle (*irm*); (iii) lateral rectus muscle (*lrm*); (iv) medial rectus muscle (*mrm*); (v) superior oblique muscle (*som*); and (vi) inferior oblique muscle (*iom*). Each of these muscles works on a particular action axis, and they are responsible for moving the eyes to the highest point of regard in the subject's field of view. According to Hering's law of equal innervation [119], the human nervous stimulus is equal and simultaneous for both eyes in a binocular visual system. Figure 44 shows the relationship between primary, secondary and tertiary positions in a typical human visual system.

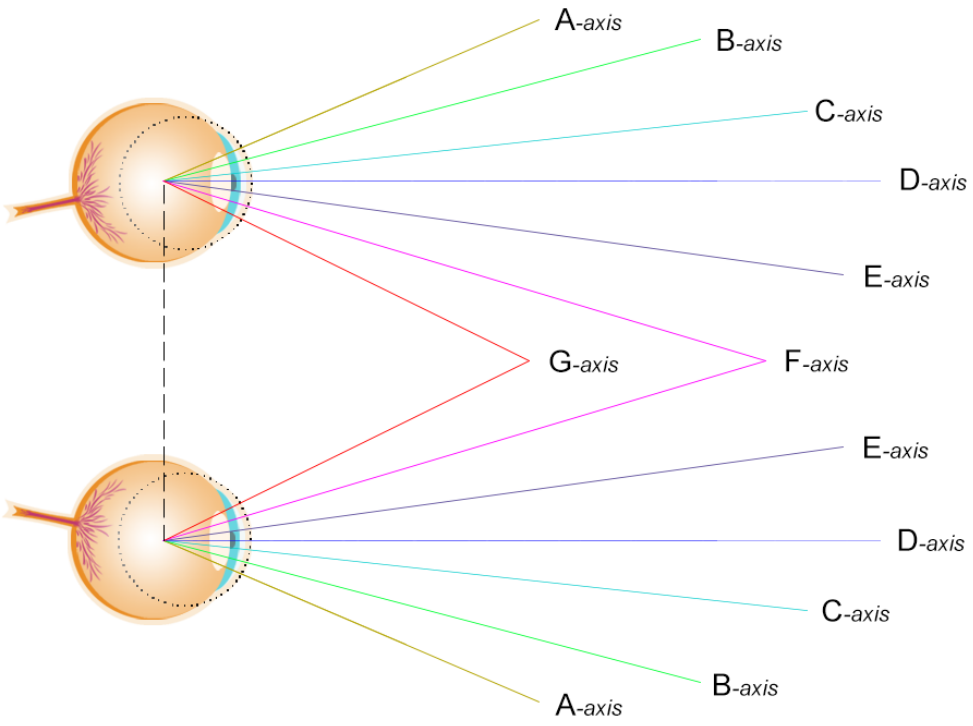


Figure 44: Human eye movements patterns in a binocular system. D-axis represents the primary position when a subject looks at the horizon. (E, F, G)-axes are the secondary positions which happen due to complex eye movements. (A, B, C)-axes are tertiary positions and are related to human rest behavior.

In Figure 44, D-axis shows the primary position in a binocular system, when a subject looks naturally at the horizon. Secondary positions (E, F, G)-axes happen due to

several eye movements, such as [120]: (i) *ductions*, ocular movements of only one eye around vertical or transversal axes; (ii) *torsions*, ocular movements of only one eye around anterior-posterior axis; (iii) *versions*, ocular movements of both eyes in the same direction, speed and amplitude which the visual axes remain in the same relative situation; and (iv) *vergences*, ocular movements of both eyes in opposite directions which the visual axes vary in their relative location. Tertiary positions (A, B, C)-axes are ocular movements corresponding to human rest (e.g. deep sleep, general anesthesia, coma, death).

Humans can move the eyes to several secondary positions around the horizontal and vertical axes when there is a visual stimulus. Eye movements of the vergence type are common sources of errors to gaze estimation methods. Vergence ocular movements can be classified as [120]: (i) *convergence*, when the initial angle (θ_1) is less than the final angle (θ_2); and (ii) *divergence*, on the contrary, as shown in Figure 45. Vergences also can be defined as *symmetric* (when the point of regard is in the middle plane of the head and angles α and β are the same) or *asymmetric* (on the contrary).

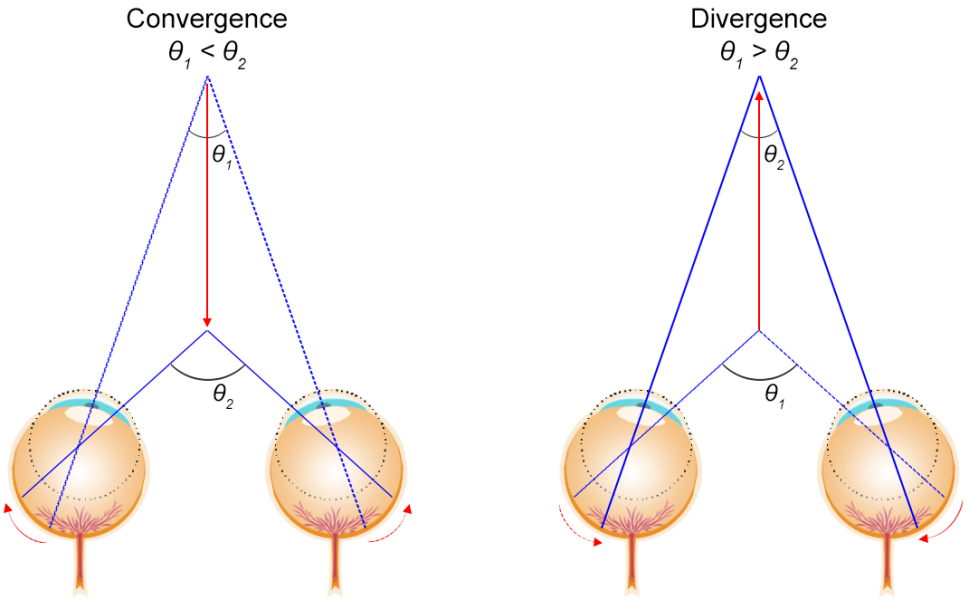


Figure 45: Vergences are ocular movements of both eyes on opposite directions which the visual axes vary in their relative situation. (*left*) convergence happens when the initial angle (θ_1) is less than the final angle (θ_2). (*right*) divergence happens when the initial angle (θ_1) is bigger than the final angle (θ_2).

6.1.1 Using the Eyes Relationship to Improve a Pupil Detector

This subsection presents a binocular eye feature detector that uses known human eye movements as priors to compensate and validate the detected eye feature. The proposed approach can be used for any eye feature, but this subsection presents only an example

based on pupil detection. In general, a binocular pupil detector aims to detected the pupil center in both eyes $[P_l(x_l, y_l),\ P_r(x_r, y_r)]$. As human eyes focus on the same object and move accordingly, it is possible to create a homography between both eyes, and use the mapping as priors to estimate the eye coordinates in two distinct ocular planes.

According to Hartley and Zisserman [110], a homography is a geometric transformation with 8 degrees of freedom, used to estimate from a point P_A in the plane A the corresponding point P_B in the plane B through the following equation: $P_B = H \times P_A$, which H is a 3×3 matrix created from 4 corresponding points in two different planes. For the proposed pupil detector, it is required to use the same eye feature data collected during the personal calibration. Thus, the gaze estimation method creates the traditional mapping from pupil centers to observed targets in the calibration plane. On the other hand, the proposed pupil detector creates a second mapping based on the left pupil center concerning the right pupil center.

Given the relationship between both eyes is prior known because the binocular movement is simultaneous, using only the left pupil center it is possible to predict the right pupil center through homography H and vice versa. Obviously, this is a simplistic model which it is not considered complex eye movements like convergence or divergence concerning the calibration plane. However, this approach can assist the eye tracking system to be more robust in the eye feature detection as well as to validate if both detected coordinates are true positive pupil candidates.

Let, $P_l(x_l, y_l)$ as the left pupil center, $P_r(x_r, y_r)$ as the right pupil center, and H_{eyes} the geometric mapping between both eyes. The proposed validation approach checks if both pupil center coordinates correspond to true positive pupil candidates, comparing the coordinated of pupil center P (detected through the pupil detector) related to the pupil center $P^{'}$ (estimated through the geometric mapping). Thus, $P_r^{'} = H_{eyes} \times P_l$ and both pupil centers will be valid if and only if $P_r \approx P_r^{'}$ (based on Euclidean distance). The same rule applies to validate the left pupil candidate.

Furthermore, the geometric mapping can remove outliers and perform a new attempt to detect the pupil when the eye feature detector fails for one of the subject's eyes. For this last case, the geometric mapping H_{eyes} is used to estimate the pupil location corresponding to $P^{'}$ and to perform a new analysis in its N-neighborhood to ensure that the detector will always detect both pupils in the processed image. Figure 46 shows the flow diagram of pupil detection developed during this project and used for sports experiments. The detector uses priors to validate and to improve the binocular pupil detection.

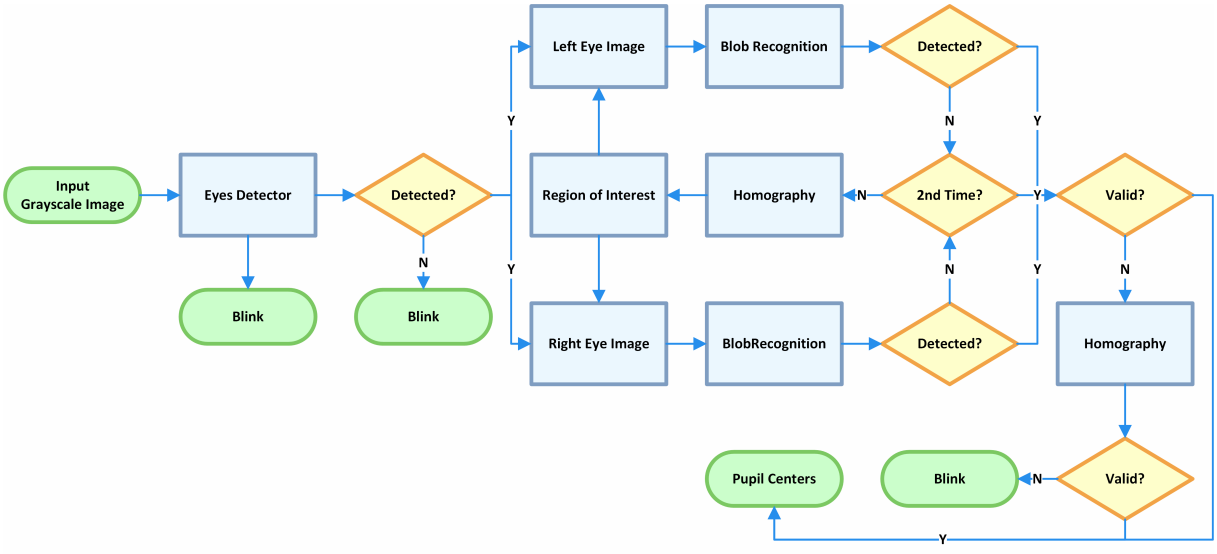


Figure 46: This flow diagram presents the proposed eye feature detector using information from both eyes to improve the robustness of pupil detection. When some information is lost, the detector uses a geometric transformation between both eyes to perform a new recognition trial. The same geometric transformation is also used to validate if both detected blobs are true positive pupil candidates.

6.1.2 Using *N*-Neighborhood to Filter a Gaze Signal

In eye tracking, gaze estimation is represented in two-dimensional gaze coordinates (g_x, g_y) on a tracking plane. However, it is possible to constitute a sequence of gaze coordinates in a one-dimensional signal, i.e. a plot with X and Y coordinates variations concerning the time t. Through the analysis of the gaze signal, the specialist can have an overview of eye movements behavior when the subject performs specific tasks, e.g. text reading [71, 121], surfing on web pages [122], understanding source-code [123], sports practices [124].

A one-dimensional signal of eye movements is a valuable data source, once it is possible to identify some gaze behaviors like blinks, fixations, saccades, glissade, smooth pursuit, microsaccades, quiet eye, among others [125]. However, the use of raw eye tracking data can include several kinds of noise from different steps of eye tracking. This subsection presents a gaze signal filtering approach that uses a range in the one-dimensional signal as priors to remove noise from the signal itself and make the eye tracking system more robust to several variations occurred during data collection and data processing.

Figure 47 shows a sequence of six video frames of a subject's left eye. The first row corresponds to the sequence of processed images, the red plot corresponds to their x_i-coordinates converted into a one-dimensional signal, and the blue plot is their y_i-coordinates. In both plots, it is possible to see an outlier in the pupil detection, which generates a pulse in both signals at time 3. In an *offline* analysis of gaze data, there is

the advantage of search for information in a N-neighborhood beyond the current frame to check if the current pupil candidate is valid or not.

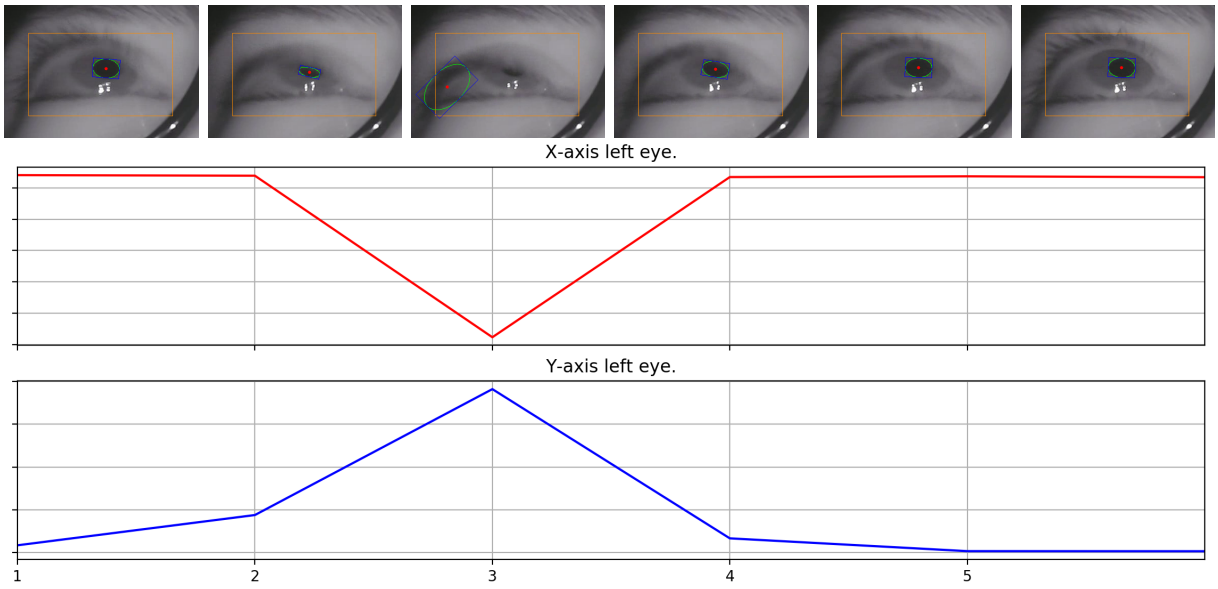


Figure 47: A sequence of video frames processed through a pupil detector based on blobs recognition. The red plot shows the x-coordinates pupil centers converted into a one-dimensional signal and the blue plot shows the y-coordinates. It is possible to see a pulse in both signals and concluded that there was a noise in the pupil detection (as shown in the third video frame).

The video sequence showed in Figure 47 was recorded using a single eye camera with 30 frames per second in a resolution of 1280×720 (*Full HD*). The illustrated gaze signal starts in a continuous region, reaches its peak in only one frame and returns to the same continuous region in the next frame, i.e. an eye movement of ± 60 ms. The proposed gaze signal filtering approach assumes that this pulse is an outlier and adjust the current pupil center (x_i and y_i) according to the coordinates of its N-neighborhood frames.

This subsection also proposes a second gaze signal filtering approach that compares the behavior of two synchronized signals created from binocular information. This approach is indicated for both *online* and *offline* gaze data analysis. In the case of only one signal presents a pulse in time t_i, the proposed approach considers the pulse as an outlier in one of the eyes. The noise can be filtered in two different ways, namely (i) through the analysis of N-neighborhood frames (*offline*); and (ii) through the mapping between both eyes to estimate the location of the missing pupil (*online*) (see Subsection 6.1.1).

Figure 48 shows the variations in x_i-coordinates of an athlete's right eye during the kayak experiment in an outdoor environment (with sudden changes of illumination). The green plot shows the signal generated from the raw pupil center and the magenta plot shows the signal filtered using a combination of both gaze signal filtering approaches.

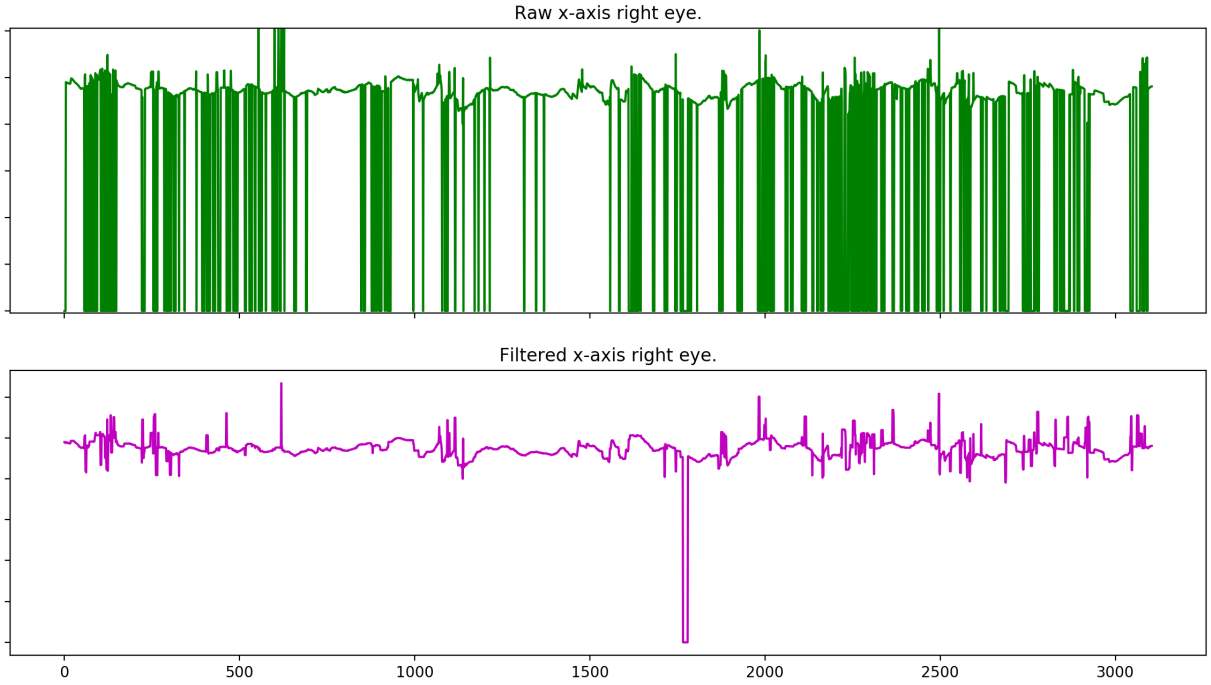


Figure 48: x_i-coordinates of pupil center converted in a one-dimensional signal. The green plot shows the raw coordinates from the athlete's right eye collected during a kayak experiment. The magenta plot shows the signal filtered based on the analysis of N-neighborhood frames and the geometric transformation between both athlete's eyes. The curve around time 1750 means one eye blink in a sequence of images.

The use of filtered signal $(P_x^{'}, P_y^{'})$ as input data to the eye tracking method ensures that the gaze estimation will be similar to the actual subject's point of regards during the eye tracking session. The gaze signal filtering approaches are linear concerning the time. Both approaches provide better results in *offline* gaze data analysis once it is possible to know the signal behavior in a wider range $n > i$ from the current processing time t_i.

6.1.3 Using Histogram Analysis in Automatic Thresholding

The proposed eye feature detector (see Figure 46) is based on blobs recognition and requires a *threshold* to convert the input grayscale image into a binary image. The binary image has different sets of pixel elements (white regions) that share between them a specific range of grayscale intensities and define eye feature candidates (pupil, iris, glints). In general, pupil detectors based on blobs recognition use classifiers to analyze some blobs properties (e.g. area, perimeter, circularity, bounding box) and define pupil candidates.

In most of viewing angles, the pupil blob appears as an ellipse and this region can be modeled by five shape parameters [17]. Fitting an ellipse to the pupil blob contours provides an accurate way to estimate the pupil center for both horizontal and vertical eye positions. However, a small change in thresholding can deform the pupil blob, decrease the

accuracy of pupil center detection and, in turn, reduce the accuracy of gaze estimation. Furthermore, in the case of outdoor scenario, the light conditions (e.g. illumination changes, shadows, reflections, different grayscale intensities for left and right eyes) may change significantly during an eye tracking session which makes it hard to set (fixed or dynamically) thresholds for pupil detection.

The grayscale intensities in the processed image (i.e. histogram) is used to define pupil thresholds automatically. For this, it is necessary to calculate the mean from eye image histogram, which the intensities are distributed uniformly in several classes (i.e. histogram bars). The mean from a histogram is calculated by multiplying each value taken by the histogram bar by the fraction of the time it occurs, through the Equation 6.1:

$$\bar{x} = \frac{1}{n_{w \times h}} \sum p_i x_i, \tag{6.1}$$

which, x_i is grayscale quantization in the processed eye image and p_i is the number of pixels with grayscale intensity x_i in the processed eye image.

The mean value by itself it is not the best threshold to be used by the proposed eye feature detector. On the other hand, it is a good starting point to understand how the processed eye image is formed. For example, corneal reflections have the highest grayscale intensities close to white (255), and their values are much bigger than the histogram mean. On the contrary, pupils are the darkest regions in the processed eye image, and they have grayscale intensities smaller than the histogram mean. It means, it is possible to use a scalar γ to assist the proposed eye feature detector in automatic select the current threshold to corresponding eye image, through the Equation 6.2:

$$threshold = \bar{x} \times \gamma, \tag{6.2}$$

which, $\gamma < 1.0$ is used to pupil detection and $\gamma \geq 1.0$ to glints detection.

A nonlinear regression defines the scalar used to calculate the threshold used in corresponding processed eye image. This regression is based on a training dataset with binocular information from 42 subjects (185 eye images) of different nationalities, genders, ages, ethnicities, with and without corrective ophthalmic lenses, in different kinds of environments, and performing distinct tasks. Figure 49 shows four different distributions based on scalars and histogram properties such as mean, mode, minimum and maximum grayscale intensities. All scalars were selected manually by a specialist to ensure the best fitting ellipse in the detected pupil candidate.

The first investigation evaluated the use of histogram mean to select the scalar γ.

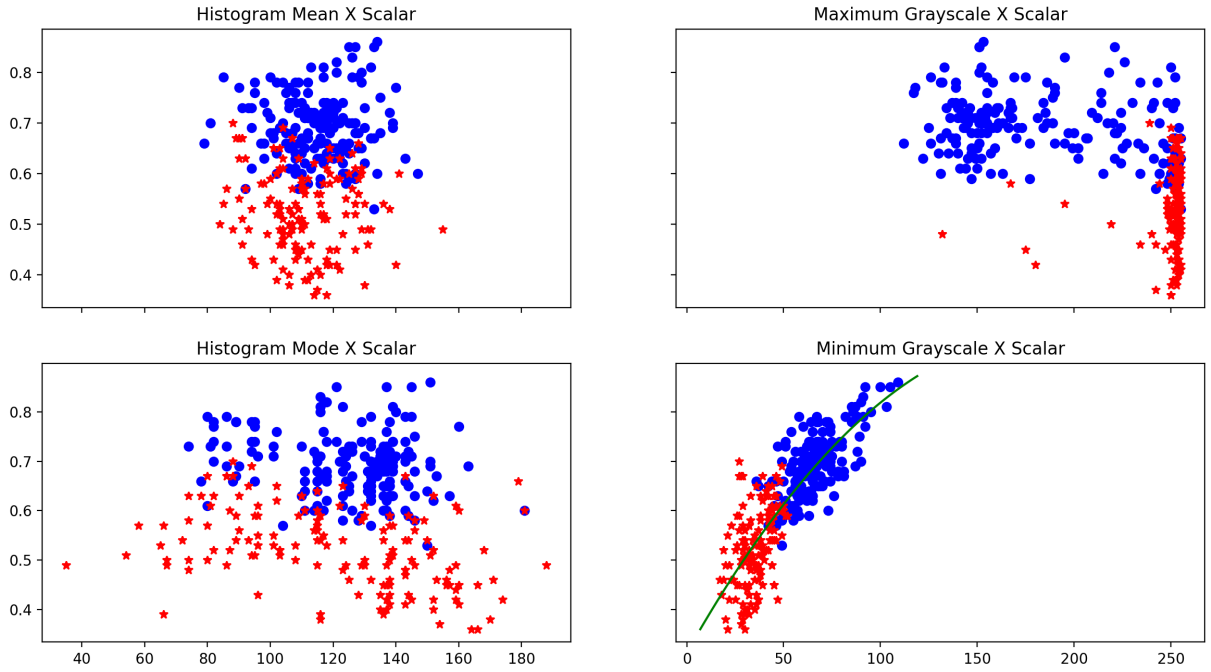


Figure 49: Threshold distribution dataset created using binocular eye images from 42 subjects. In total, 329 eye images was analyzed in different illumination conditions. The blue circles represent data collected in outdoor experiments and the red stars are the data collected in indoor experiments. (*top-left*) histogram mean vs. scalars; (*top-right*) maximum grayscale intentisities vs. scalars; (*bottom-left*) histogram mode vs. scalars; and (*bottom-right*) minimum grayscale intensities vs. scalars.

However, the data distribution concerning histogram mean is not a good dataset to create a regression model, because there is a strong correlation between data from indoor and outdoor environments. For this reason, the experiment was expanded to investigate the relationship between scalars and others histogram properties. Histogram mean (57.00 ± 57.06), mode (61.94 ± 63.91) and median (58.21 ± 58.49) presented practically the same data distribution. The largest variance was observed in data distribution using the maximum grayscale intensities from the processed eye images (106.42 ± 110.93), as shown in Figure 49 (*top-right*).

The best nonlinear regression to predict the scalar γ used to calculate the dynamic threshold is based on the minimum grayscale intensities in the processed images. The minimum grayscale intensity property presented the lowest variation among the analyzed data distributions (26.46 ± 29.09). The nonlinear regression uses a second order polynomial to fit the dataset in Figure 49 (*bottom-right*). In summation notation this regression polynomial is defined in Equation 6.3:

$$y = a_0 + a_1 x + a_2 x^2, \tag{6.3}$$

which, the coefficients in used dataset are: $a_0 = 0.312523682$, $a_1 = 0.006937170$ and

$a_2 = -0.000018764$ (see green line in Figure 49 (*bottom-right*)). After selecting the scalar γ, the pupil detector performs the blobs recognition module in a range of $\gamma \pm 0.1$ in steps of 0.01 to choose the best pupil candidate. In this case, the "*best*" pupil candidate is the one that has the best circularity (the closest value to 1), according to Equation 6.4:

$$circularity = \left| \sqrt{\frac{4 \times \pi \times area}{perimeter^2}} \right|. \tag{6.4}$$

All plots in Figure 49 show conflicts in small regions between data collected in indoor (*red stars*) and outdoor (*blue circles*) environments. The conflict happens because in some images from outdoor environments was captured in place with little influence of sun light. For example, Figure 50 (*left*) shows a kayak's athlete passing under a bridge and Figure 50 (*right*) shows that eye image has a grayscale intensity similar to the one captured in indoor environments. In general, scalars bigger or equal to 0.6 are used to define the threshold used in outdoor images, and scalars less than 0.6 are used to indoor images.

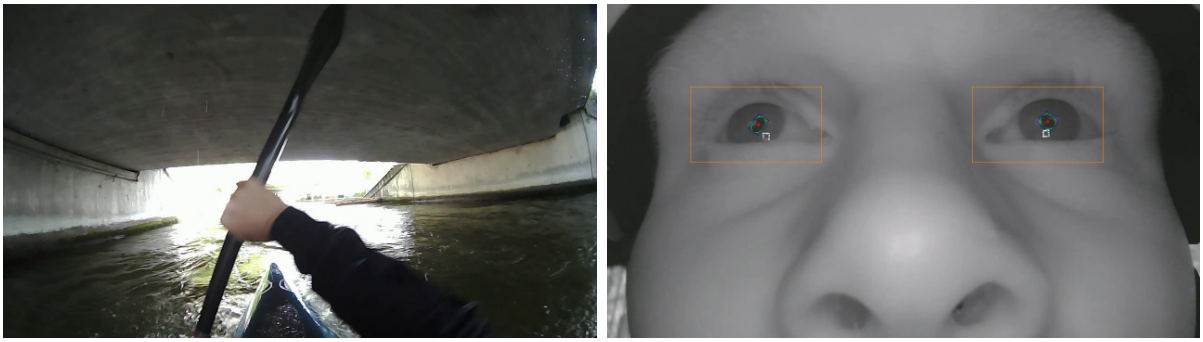


Figure 50: An example of outdoor illumination influence in eye image. (*right*) The kayak athlete passes under a bridge where there is a small influence of sun light in the captured images. (*left*) The captured eye image has similar contrast compared to the images captured in indoor environments.

Figure 51 shows the result of automatic scalar selection using the nonlinear regression polynomial (Equation 6.3) presented in this subsection. This gaze signal represents a sequence of eye images from one athlete during the kayak experiment (see Section 7.3). In this case, at several times the scalar signal exceeds the zone defined for the indoor images. These are situations where eye images were captured in places with little sun light illumination during the experiment, e.g. between the time range 1300-1400, it is possible to observe what happens when the athlete pass under the bridge. The others negative peaks are related to trees shadows, tends over the river, and the action of rotating around (*upstream*) one of the poles in kayak gates (see Subsection 7.3.1).

Eye images recorded in indoor environments suffer less influence of sudden changes

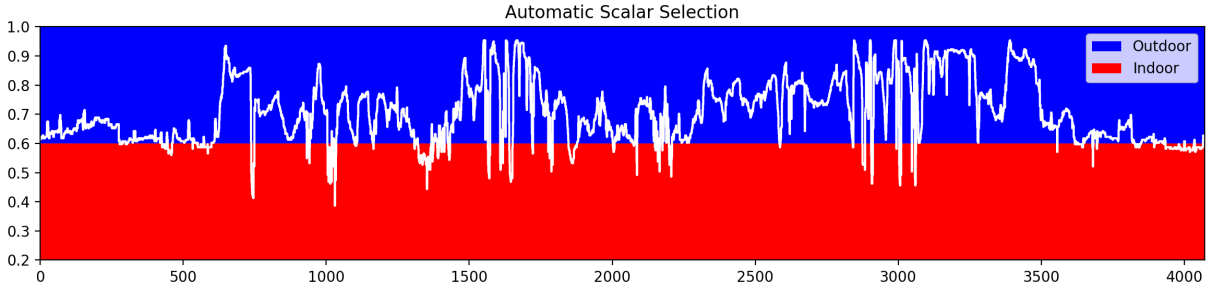


Figure 51: The result of automatic scalar selection using an eye video from kayak experiment. The signal passes through the red area (for indoor images) a couple of times because the athlete was in areas without the influence of sun light (e.g. bridges, trees shadows, covered areas).

in the threshold selection due to a better controlling of indoor illumination. Figure 52 shows a sequence of eye images captured in a laboratory experiment. In this case, at any moment the signal exceeds the outdoor zone (i.e. blue area). It is also possible to conclude that changes in the threshold selection are strongly related to eye rotations. At the beginning of eye tracking session, the participant looks at the horizon. Between time range 300-1450 the personal calibration is performed. From time 1450 onwards, the participant only follows a target moving in depth. The jumps in this interval represent some participant's blinks during the experiment.

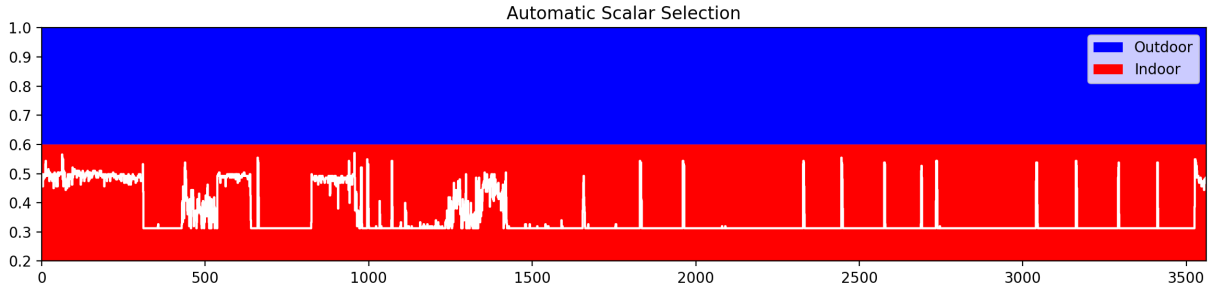


Figure 52: The result of automatic scalar selection using an eye video from a laboratory experiment. The signal never passes through the blue area (for outdoor images). Indoor images are the easiest to select the threshold thought the presented nonlinear regression automatically.

6.2 N-Closest Glint Normalization Approach

Glint normalization is a method that uses homographies to make an eye tracking system more robust to head movements in remote eye trackers [65]. However, this approach can also be used by head-mounted eye trackers [27] to reduce the influence of eye camera slippage concerning the user's head. The eye tracker may move during use e.g. when the athlete is shooting due to the impact produced by the gun. A small displacement in the

eye camera position reduces a lot the accuracy and precision of gaze estimation methods.

Hansen et al. [70] introduced a glint normalization method that selects the best degraded homographies according to the number of detected glints, namely (i) *similarity*, a transformation that preserves the shape (for 2 glints); (ii) *affine*, a non-singular linear transformation that preserves the length ratios and angles between lines because its non-isotropic scaling (for 3 glints); and (iii) *homography*, a general non-singular linear transformation that preserves the collinearity of transformed points (for 4 or more glints).

This section presents an extension of Hansen et al.'s method [65, 70], which analyzes the N-closest corneal reflections ($1 \leq N \leq 3$) from the pupil center given all true positive glints. The proposed glint normalization approach uses all geometric transformations from Hansen et al.'s method [70] and the *Euclidean distance* for a single detected glint.

The simplified human eye geometric model assumes that pupil and corneal reflections coincide on corneal plane. However, corneal reflections are non-linearly distorted on corneal curvature surface and the planarity assumption can enhance the gaze estimation error. Using the N-closest glints approach, it is possible to investigate the influence of non-linearly distribution of glints on the cornea and compare it with the traditional homography normalization method [65, 70].

6.2.1 Assessment using Simulated Gaze Data

The evaluation of proposed glint normalization approach used simulated gaze data in a controlled environment for MATLAB [113], where it is possible to evaluate the effects of noise in each eye parameter. The data analysis was based on both multiple geometric transformations and N-closest approach. The assessment was divided in two main tasks (i) the observed targets on the screen plane with the head still, and (ii) a single observed fixed target in different head positions. The evaluation investigated the error distribution in gaze estimation concerning: (1) refractive indexes of aqueous humor and cornea [α]; (2) horizontal [γ] and vertical [β] angle offset between optical and visual axes [*angle kappa*]; (3) number of calibration targets [N]; and (4) XY-axes and XZ-axes head movements.

6.2.1.1 Apparatus

The simulated eye tracker was a monocular remote eye tracker with four infrared light sources. The remote eye camera was at the center and slightly below of the screen, the infrared light sources were at the screen corners, and the screen was 55 cm away

from the user. Each experiment analyzed the gaze estimation from 256 observed targets distributed in a 16×16 matrix over the screen plane. All tests aimed to investigate the gaze estimation from two different eye models [65] with distinct angle kappa offsets, namely (i) E_0 ($\gamma = \beta = 0°$), and (ii) E_1 ($\gamma = 4.5°, \beta = 1.5°$). Although, the eye model E_0 is physically infeasible setup, it assists to avoid some eye biases into the controlled environment. For standard, the experiment used the minimum number of calibration targets ($N = 4$) necessary to calibrate a gaze estimation method based on homography [65].

6.2.1.2 Evaluation Controlling Head Movements

Refractive Indexes of Aqueous Humor and Cornea The test evaluated the influence of eye refractions to the glint normalization, as shown in Figures 53 and 54. According to Hansen and Ji. [17] the refractive index of aqueous humor is around 1.336, and the refractive index of the cornea is around 1.376. The results of this test showed that the refractive indexes present only a slight influence in the glint normalization.

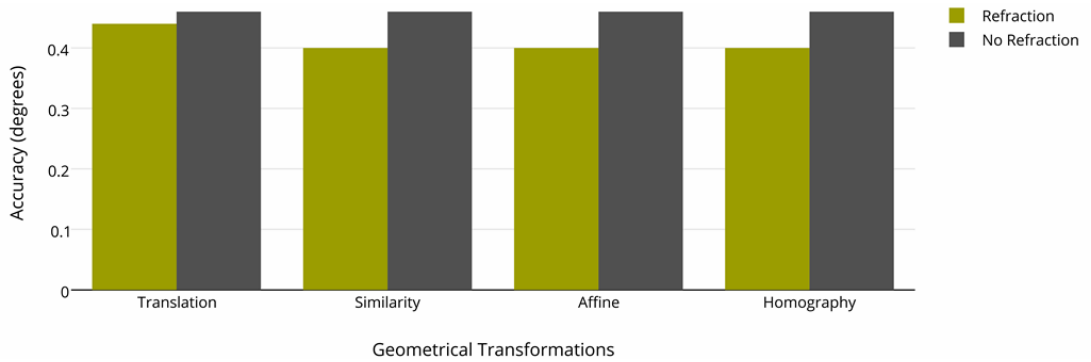


Figure 53: The influence of the refractive index of aqueous humor [1.336] and the refractive index of cornea [1.376] to proposed glint normalization using the eye model E_0 [$\gamma = \beta = 0°$].

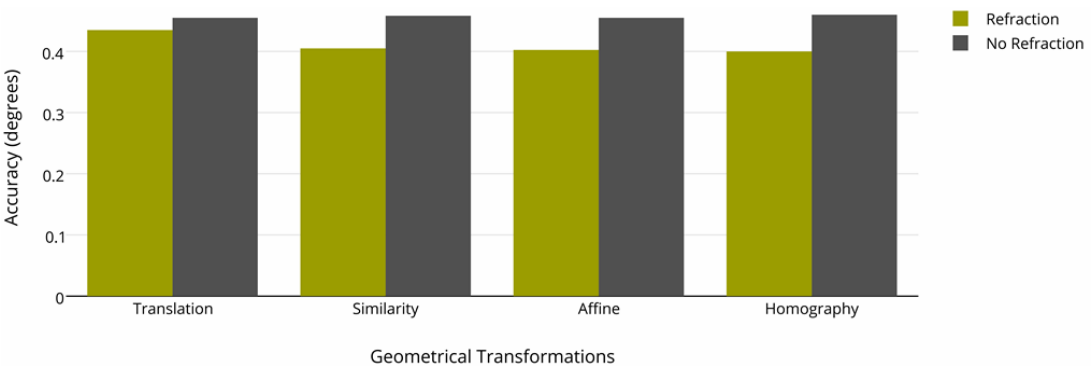


Figure 54: The influence of the refractive index of aqueous humor [1.336] and the refractive index of cornea [1.376] to proposed glint normalization using the eye model E_1 [$\beta = 1.5°, \gamma = 4.5°$].

Optical and Visual Angle Offset This test evaluated the influence of different angles kappa to the glint normalization. Figure 55 shows the accuracy as a function of various horizontal angles ($-4.5° \leq \gamma \leq 4.5°$) and a fixed vertical angle ($\beta = 0°$). The results of this test showed that all geometric transformations from Hansen et al.'s method [70] model with high accuracy (i.e. $error < 0.5°$) the angle kappa during the personal calibration.

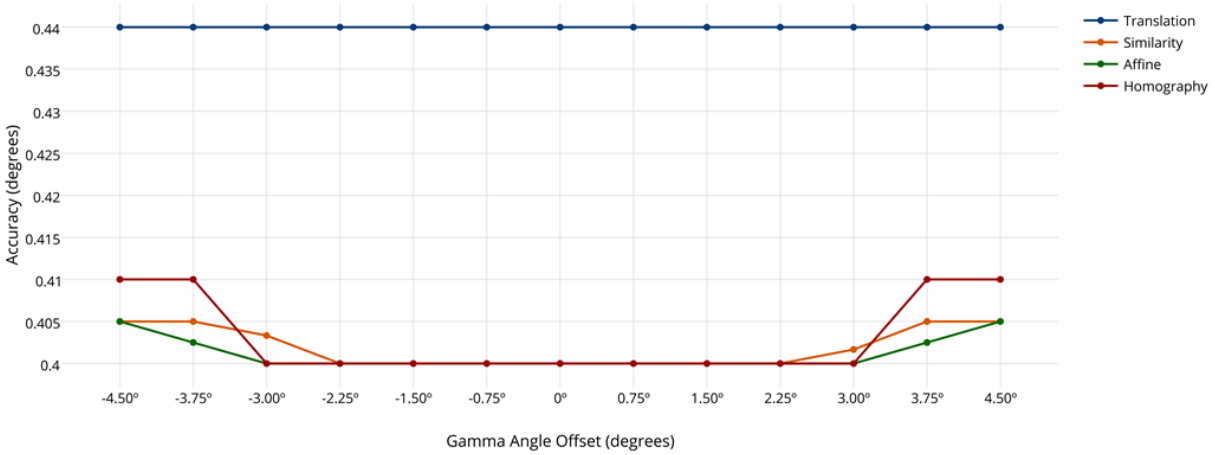


Figure 55: The influence of horizontal [γ] and vertical [β] angle kappa offset to proposed glint normalization. This test investigated the influence of angle kappa with $-4.5° \leq \gamma \leq 4.5°$ and $\beta = 0°$.

Number of Calibration Targets This test investigated the influence of the number of calibration targets to the glint normalization based on the following range: $4 \leq N \leq 25$. Figures 56 and 57 show the accuracy of gaze estimation as a function of the number of calibration targets for both E_0 and E_1 eye models. The evaluation of similarity, affine transformation, and homography presented similar results. On the other hand, Euclidean distance presented a similar performance curve, but a worse accuracy (around $+0.05°$).

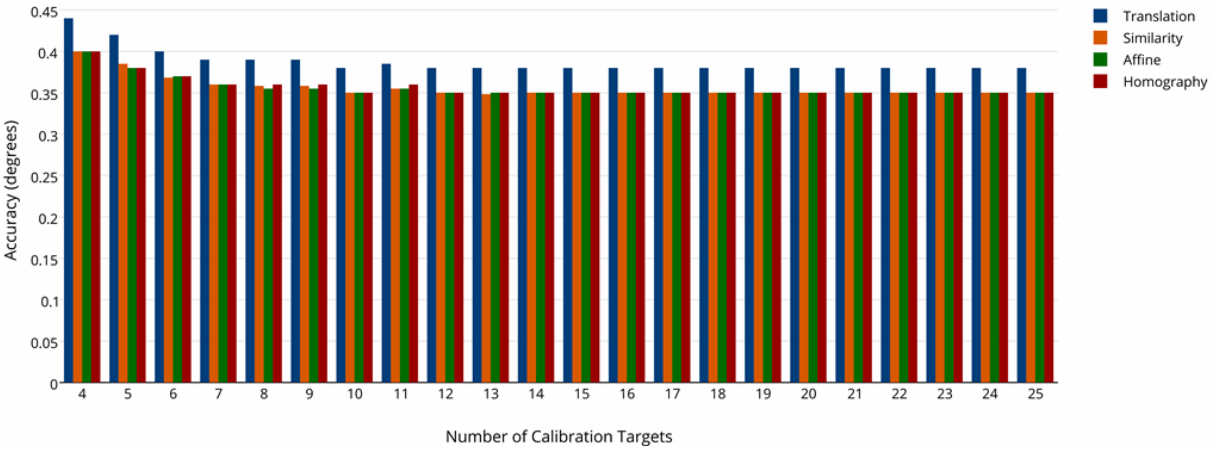


Figure 56: The influences of the number of calibration targets ($4 \leq N \leq 25$) to the glint normalization using the eye model E_0 [$\gamma = \beta = 0°$].

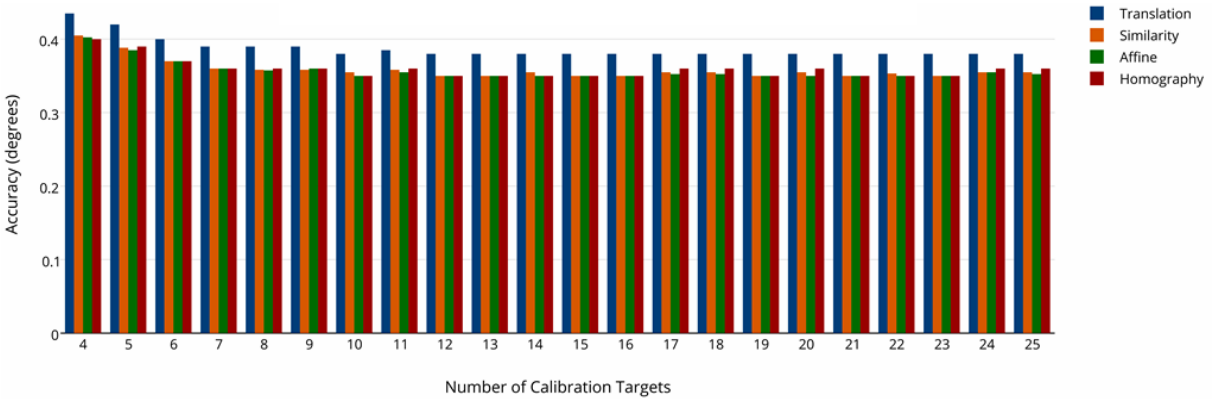


Figure 57: The influences of the number of calibration targets ($4 \leq N \leq 25$) to the glint normalization using the eye model E_1 [$\beta = 1.5°, \gamma = 4.5°$].

6.2.1.3 Evaluation with Head Movements

Head Movements Around XY-axes During the experiments to investigate the robustness of glint normalization to the head movements around XY-axes, the eye looked at a fixed point on the screen while the head moved 16 positions on the X-axis and 16 positions on the Y-axis. The fixed point was one of the calibration targets because gaze estimation on this point presents high accuracy [65]. This experiment evaluated the robustness of the proposed glint normalization approach with N-closest glints concerning the head movements as well. Given four corneal reflections, the proposed approach uses the 3-closest glints for affine transformation, 2-closest glints for similarity, and 1-closest glint for Euclidean distance. Figure 58 shows that the 3-closest and 2-closest normalizations present a performance similar to traditional homography normalization [65].

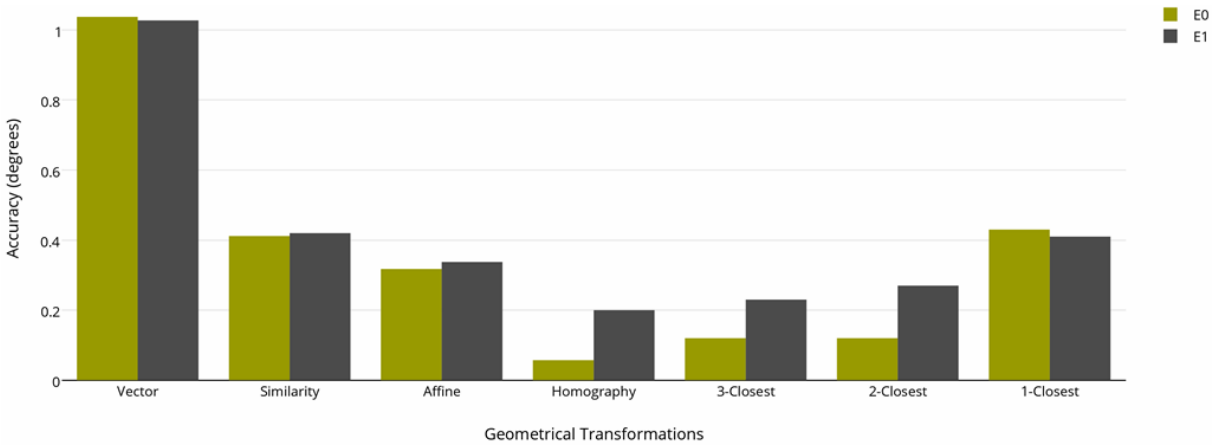


Figure 58: The influences of head movements around XY-axes to proposed glint normalization approach using eye model E_0 [$\gamma = \beta = 0°$] and eye model E_1 [$\beta = 1.5°, \gamma = 4.5°$]. The eye location changes in steps of 100 mm in both x- and y-coordinates from the camera center.

Head Movements in Depth The last simulated test evaluated the robustness of proposed glint normalization approach to head movements in depth. Again, a fixed target remained still on the screen while the eye moved in steps of 100 mm in X-coordinate and 300 mm in depth. During the personal calibration, the screen as at 550 mm away from the user. For the experiment, the eye moved in a range of 400-700 mm in 16 positions. The 3-closest and 2-closest normalizations presented a better performance for head movements in depth compare to traditional homography normalization [65], as shown in Figure 59.

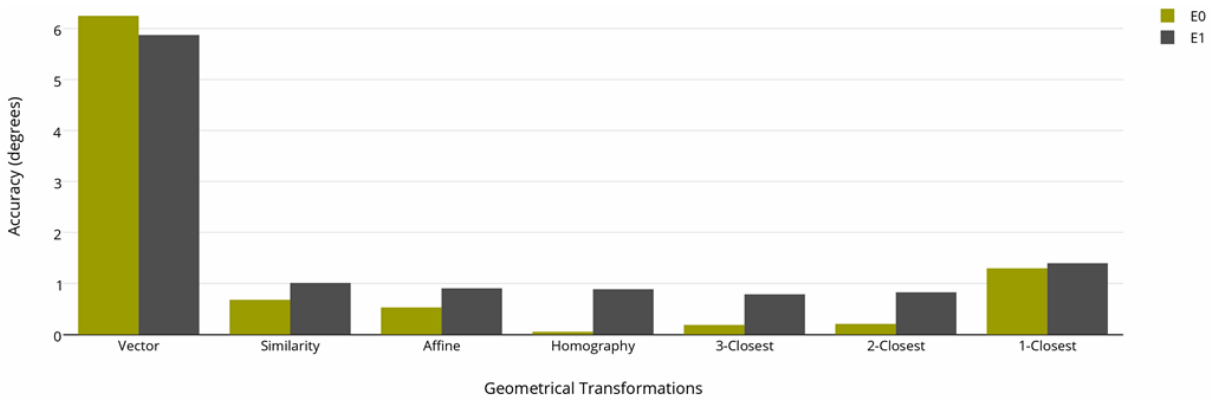


Figure 59: The influences of head movements in depth to proposed glint normalization approach using eye model E_0 $[\gamma = \beta = 0°]$ and eye model E_1 $[\beta = 1.5°, \gamma = 4.5°]$. The eye location changes in steps of 100 mm in X-coordinate and 300 mm in depth.

6.3 Eye Camera Slippage Compensation Model

Although glint normalization approach proposed in Section 6.2 works very well for both remote and head-mounted eye tracking applications, some practical problems arose during sports experiments. In general, infrared light sources are attached close to eye cameras in head-mounted eye trackers. However, eye cameras are located slightly below to subject's eyes and corneal reflections are not present in the most of the captured eye images, mainly due to large eyeball rotations.

It is necessary to use any reference point to compensate the eye camera slippage. This section presents a new eye feature normalization approach to be used in missing corneal reflections. Eye corners are features widely used to eye feature normalization and to assist gaze estimation [40, 57, 58, 126]. Even after small displacements of eye cameras around x- or y-coordinates concerning the user's head, the relationship between eye corners and pupil center do not suffer major changes. In general, eye corner detection algorithms are unstable, and they add noise to eye feature normalization and to gaze estimation.

6.3.1 Eye Feature Normalization Approach

This subsection presents a novel eye feature normalization approach that uses landmarks around the eyelids to normalize the pupil center. The proposed approach created a particular labeled binocular landmark dataset using 260 infrared eye images. The eye images were manually labeled, specifying eight (x, y)-coordinates of landmarks around each eye, and the two regions of interest around the eye regions. The training dataset was trained using a regression tree ensemble algorithm developed by Kazemi and Sullivan [127].

The binocular landmark positions are directly estimated from the grayscale intensities themselves. The proposed approach improves the controlling of eye camera slippage in head-mounted eye trackers and, at the same time, maintains the robustness of eye feature normalization during the entire eye tracking session. Figure 60 presents four examples of eye images used to test the learning approach. The processed images contain eight landmarks around each eye contour (i.e. green points) and two regions of interested with recognized eyes (i.e. red rectangles) in a binocular image.

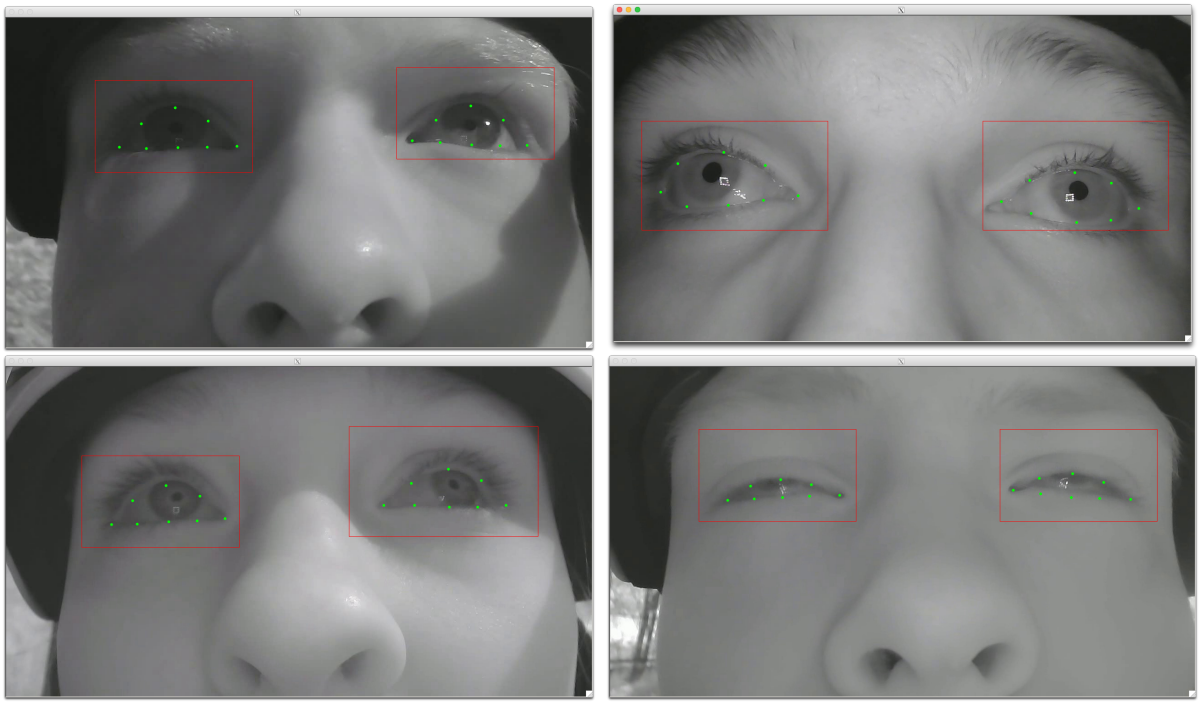


Figure 60: An example of four eye images used to test the performance of a learning approach. It can recognize both eyes in the processed image (red boxes) and, at the same time, the eye contour through pre-defined landmarks (green points).

Each landmark contains one ID_i in a range of $0 \leq i \leq 7$, which ID_0 are lateral eye corners, ID_2 are top landmarks, ID_4 are nasal eye corners, and ID_6 the bottom landmarks for each eye. Only landmarks ID_0 and ID_4, and the midpoint of the segment

connecting both landmarks are used in eye feature normalization because they do not change their locations during eyelids movements, as shown in Figure 60 (*bottom-right*). The normalization is based on Euclidean distance between the pupil center and the closest reference point (i.e. ID_0, ID_4, or the midpoint). The closest approach is used to get better performance in the normalization as shown in Subsection 6.2.1.3.

The same algorithm is used to detect the user's eyes and define regions of interest in the processed image, as presented in the flow diagram in Figure 46. Thus, the pupil detection is performed only on each region of interest, and it is possible to speed up the processing time and reduce the number of false positive pupil candidates. The algorithm also has a routine to detect eye blinks (see the first blink output in Figure 46) based on the aspect ratio of landmarks ID_0, ID_2, ID_4 and ID_6, using the Equation 6.5:

$$ratio = \frac{distance_h}{2 \times distance_v}, \tag{6.5}$$

which $distance_h$ is Euclidean distance between landmarks ID_0 and ID_4, and $distance_v$ is Euclidean distance between landmarks ID_2 and ID_6. The algorithm sets the currently processed eye image as an eye blink when $ratio < 0.09$ and stops the eye feature detection.

6.3.2 Assessment using Real Gaze Data

The evaluation of the proposed eye features normalization approach used a real eye tracking scenario in a laboratory experiment. The participants used the same head-mounted eye tracker built for the head-rotation experiment (see Subsection 5.3.1). A sample of 5 volunteer participants (4 males and 1 female) were recruited from the IT University of Copenhagen. Participants ranged from 26 to 45 years old ($mean = 34.2 \pm 6.9$) and their height ranged from 151cm to 187cm ($mean$ = 173.8cm ± 13.7cm). Two of them had normal vision, two wore glasses, and one wore contact lenses.

During the personal calibration, the participants stood in a fixed position and looked at 9 targets in a whiteboard at 4 meters from them. Participants were instructed to look at the calibration target number 5, keep their focus of attention at the target during the entire experiment and rotate the head freely for how long they wanted. At the end of the experiment, the participants performed a recalibration to compare the error distribution concerning the calibration data collected at the beginning of the eye tracking session. Participants were free to blink, move the body and withdraw from testing at any stage.

Figure 61 shows the (x, y)-coordinates distribution of a participant's right eye during

the execution of experiment. It is possible to see in Figure 61 (*bottom*) that the plot has a descendant curve when comparing the beginning and the end of eye tracking session. The difference between initial and final Y-coordinates in the blue plot is around 65 pixels.

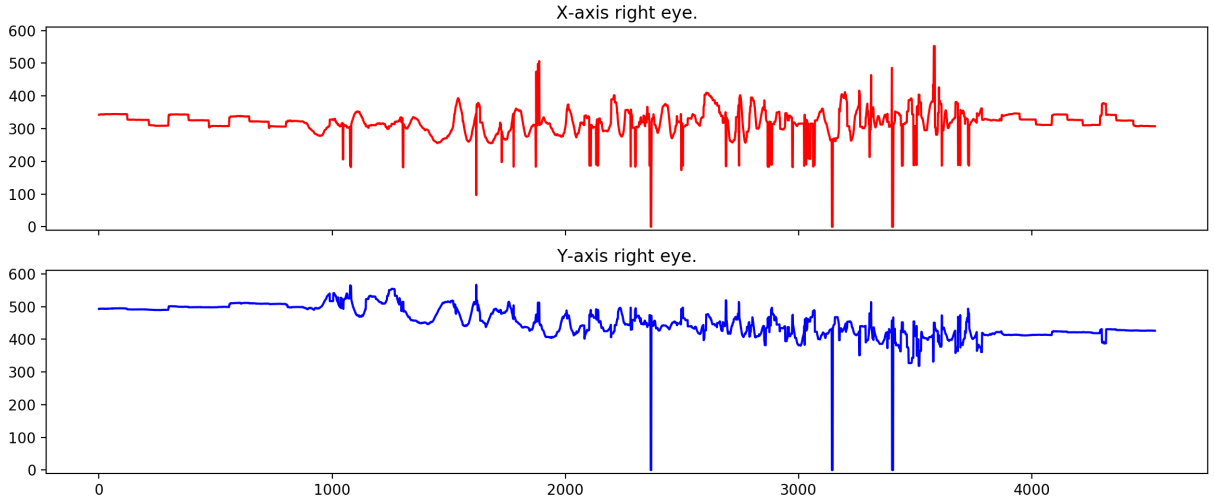


Figure 61: Influence of eye camera slippage in the eye feature detection. (*top*) X-coordinates do not change along the eye tracking session, while (*bottom*) shows a descendant curve in Y-coordinates.

Eye corners are used to identify the eye camera slippage. The compensation approach (see Subsection 6.3.1) uses eye corners as reference points in an eye feature normalization. The biggest problem found in this experiment was the instability of binocular landmark detector because both eye corners (inner and outer) presented considerable variations along of processed eye video. Figure 62 shows Y-coordinates from the inner right eye corner of the same participant's eye illustrated in Figure 61. The blue plot in Figure 62 shows several jumps with large dispersion along the entire eye tracking session.

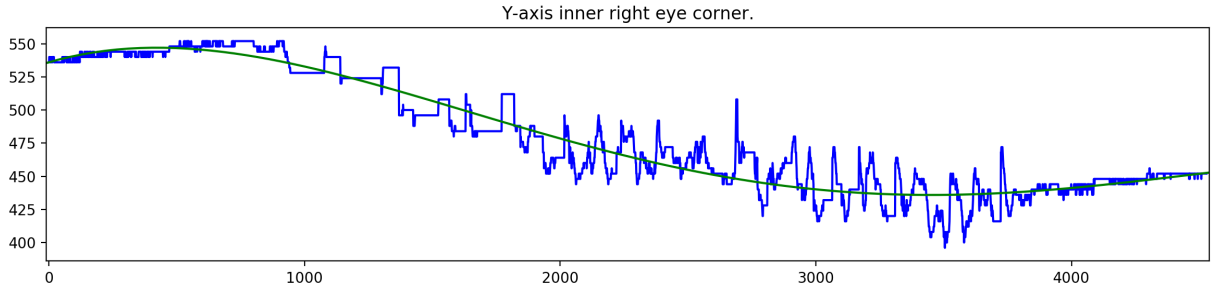


Figure 62: (*blue*) Y-coordinates from inner right eye corner along the experiment. The blue plot presents a descendant curve related to the eye camera slippage concerning participant's head. (*green*) A 4^{th} order polynomial regression to calculate the eye corner coordinate used into eye feature normalization.

Figure 61 (*bottom*) and Figure 62 show that pupil center and eye corners coordinates presented similar descendant curve in Y-coordinates along the time. A fourth order

polynomial regression fits the distribution of the eye corners coordinates (see the green plot in Figure 62) to reduce outliers in the eye feature normalization. This polynomial regression provides the reference points used in the eye feature normalization approach.

Figure 63 shows a plot with the average gaze error of nine targets observed during the recalibration, i.e. the difference of gaze estimation between the beginning and the end of the experiment. The red bars are the gaze error distribution using a gaze estimation method based on homography [65], and the green bars are the gaze error distribution using the proposed eye feature normalization approach before estimating the gaze.

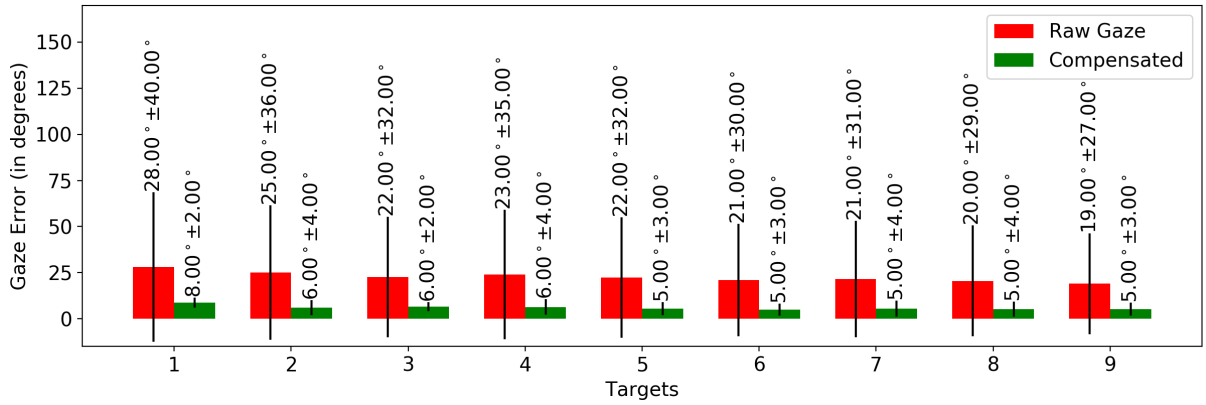


Figure 63: Average gaze error between the beginning and the end of an eye tracking session. (*red bars*) Average gaze error using a traditional gaze estimation method based on homography [65] and using only the pupil center as input data. (*green bars*) Average gaze error of proposed approach that uses normalized pupil center as input data to the same gaze estimation method based on homography [65].

Figure 63 shows the average gaze error distribution over the 5 participants. In the first analyzed item (i.e. the red bars), the pupil center is the only one input data used to the gaze estimation method, which the mean gaze error is $22.73° \pm 33.14°$. The significant variance is related to the eye camera slippage. In the second analyzed item (i.e. the green bars), the proposed approach normalizes the pupil center, and then the normalized pupil is the input data used to the gaze estimation method, which the mean gaze error is $5.99° \pm 3.57°$. The proposed eye feature normalization approach presented an improvement of 74.94% in this dataset.

The proposed eye feature normalization approach showed very promisingly to reduce the influence of eye camera slippage in uncalibrated head-mounted eye trackers. However, the development of proposed approach still needs many improvements to increase the robustness of eye corners detector. Thus, it is proposed to use a broader dataset with eye images from different subjects, in indoor and outdoor environments, while performing various daily tasks using a head-mounted eye tracker. It is also important to increase the

number of landmarks around the eyelids to stabilize the landmarks that represent the eye corners (i.e. currently the landmarks ID_0 and ID_4).

6.4 Conclusions

It is tough to identify the pupil and glints in eye images for outdoor applications because of illumination noise, reflections, sunlight, shadows, among others. For this reason, this chapter presents a collection of eye tracking methods to improve the robustness of a head-mounted eye tracker in the wild. These methods use eye information to reduce the influence of noise for eye feature detection. The use of binocular information to validate and compensate the pupil center showed very promisingly in outdoor and indoor experiments, in particular for sports practices (kayak and shooting). Obviously, the mathematical model based on homography is simplistic, and it does not take into account that after the personal calibration complex eye movements could happen which the mapping is not able to handle, e.g. ductions, torsions, versions. However, some situations which one of the eyes suffer a strong influence of external noises (e.g. shadow, sunlight), the proposed binocular compensation approach presented satisfactory performance.

The offline analysis of gaze data allows searching information in different ranges of the gaze signal and, at the same time, having an overview of gaze behavior during the entire eye tracking session. For example, the pupil center coordinates (pc_x, px_y) can be represented in two one-dimensional gaze signals for each X- and Y-coordinates. The analysis of N-neighborhood range allows assisting filtering the input gaze data before the gaze estimation. However, it is important to be aware that the proposed gaze signal filtering approach works very well for eye tracking application with normal frame rate (30-90 Hz). It would have a different behavior to remove outliers from high-speed gaze signal (e.g. 1000 fps). The offline analysis also allows using gaze signals from both eyes to remove outliers and to identify gaze behaviors like blinks, saccades, and fixations. This approach was widely used during this project to process the eye tracking data collected from eye tracking experiments in real scenarios.

Another contribution presented in this chapter was a learning approach to assist the eye feature detector in recognizing the environment setup (indoor or outdoor) and selecting the best threshold for binarizing the input eye image. This learning approach uses the minimum grayscale intensity of processed eye image to choose the scalar used to calculate the current threshold. This approach presented an excellent performance for

the entire dataset created in this project, but it was not validated to eye feature dataset in general. Before publishing the proposed automatic threshold selection, it is necessary to validate it using others pupil datasets (e.g. [33, 34, 128]).

This chapter also presents a novel glint normalization approach based on the extension of multiple geometric transformations [70]. This glint normalization approach is robust to noise and selects the best geometric transformation according to the number of detected corneal reflections. The primary intention was creating a robust glint normalization that works even with one glint. However, the first experiments showed that translation transformation is intolerant to head movements. For this reason, the proposed approach uses the Euclidean distance between pupil center and detected glint as an alternative to translation normalization. During the personal calibration, it is necessary to create all geometric transformation matrices, i.e. four matrices for the norm, six matrices for similarity, four matrices for affine transformation and one matrix for homography normalization. The biggest contribution of the proposed approach is the N-closest glints normalization. This normalization presents a performance similar to traditional homography normalization [65] even using only two corneal reflections. The proposed approach will be submitted as a paper to some conference in the field of Computer Science and its draft is available in Annex C.

The normalization is useful to compensate the influence of eye camera slippages concerning the subject's face along the eye tracking session. However, glint normalization presents some practical problems for head-mounted eye trackers because glints are not available in all processed eye images due to large eyeball rotations. For this reason, this chapter presents an alternative eye feature normalization based on eye contour landmark detection. The eye corners coordinates are used to normalize the pupil center before the gaze estimation and it presented an improvement of 74.94%. However, the current eye contour landmark detector based on regression tree ensemble algorithm was trained using a small dataset and only eight landmarks around each eye. Thus, eye corners detection presented significant variation in (x, y)-coordinates as shown in Figure 62. A fourth order polynomial regression was used to reduce the outliers of eye corners coordinates and to calculate the reference point used to normalize the input pupil center.

The proposed eye feature normalization is begin improved in collaboration with the University of the West of Santa Catarina (UNOESC), supervised by Prof. Lilian Jeannette Meyer Riveros and the Bachelor student in Computer Science Fernando Wolff. In this new study, the learning approach will be trained and tested using a dataset with

binocular information from 42 subjects (185 eye images) of different nationalities, genders, ages, ethnicities, with and without corrective ophthalmic lenses, in different kinds of environments, and performing distinct tasks (i.e. the same dataset used to training the automatic threshold selection in Subsection 6.1.3). The new study proposes to use 64 landmarks around the eyelids to increase the stability of eye corners detector. This eye feature normalization approach will be submitted as a full paper to 2018 Symposium on Eye Tracking Research & Applications (ETRA 2018).

Part III

Eye Tracking in Sports Experiments

Chapter 7

Experiments in Sports Scenarios

This chapter presents two eye tracking experiments performed in sports scenarios. The first experiment planned with the objective to explore different combinations of hardware to improve the eye tracker for the chosen sports. The second one with the aim to collect and analyze gaze data in the wild to support the development of the eye tracking methods proposed in this Ph.D. thesis, which are robust to critical situations in indoor and outdoor environments.

The dataset generated through these experiments and laboratory experiments contains eye information from more than 100 distinct subjects of different nationalities, genders, ages, ethnicities, with and without corrective glasses, in various environments, and performing distinct sports tasks. In a scope of the following experiments, an elected group of that dataset have been used in different moments.

Section 7.2 describes the shooting experiments, which assisted in developing an eye tracking system and in building some head-mounted eye tracker prototypes. A collaboration between the IT University of Copenhagen (Denmark) and the University of Leipzig (Germany) resulted in a kayak experiment as described in Section 7.3.

7.1 Collaborations

This Ph.D. thesis had an extensive collaboration with the *Danish Elite Sport Organization* (aka Team Danmark), which provided access to the training facilities, athletes, and experts during the shooting experiments. Eye tracking methods and tools developed during this Ph.D. thesis were used at the *Department of Movement and Training Science of Natural Sports* of the *University of Leipzig*, in a series of experiments coordinated by the Ph.D. students *Melanie Mack* and *Otto-Max Klein*. There was also a collaboration with the *Deutscher Kälteund Klimatechnischer Verein* (DKV), during the kayak experiments.

7.2 Shooting Experiment

The shooting experiment was an environment for continuous tests, controls, and appraisal of the eye tracking methods proposed in this thesis, and to test and found the best combination concerning the eye tracker hardware. It worked as a laboratory test, where for each new software or hardware update released, a new daily training session was scheduled with a shooting athlete to investigate the need for new enhancements in the hardware, and to identify bugs in the eye tracking system methods in development. The idea of creating adaptable eye trackers that exploit the constraints revealed for specific sport training settings has emerged during shooting experiments. Each shooting experiment session used the same methodology protocol with a single one elite athlete.

7.2.1 Method

Participants The elite athletes who participated in most of shooting experiments were (i) *Mikkel Petersen*, junior elite athlete with 8 years of experience, Danish champion in his category in 2016; and (ii) *Jesper Hansen*, expert elite athlete with 24 years of experience, world champion in 2014 and fifth position Rio 2016 Olympic Games. Figure 64 shows an athlete during a daily training session using one of our head-mounted eye trackers.

Figure 64: Mikkel Petersen's eye movements are collected while the target is rushing through the air, on the skeet field at Københavns Flugtskytte Klub.

Apparatus In a case of shooting, according to Danish laws athletes are obliged to use

protection glasses and hearing protection in the daily training sessions and official competitions. Consequently, the head-mounted eye trackers prototypes used in the shooting experiment were built using frame glasses and, in some cases, helmets. The set of equipment had constantly changed, during de experiments as shown in Figure 65. The essential configuration included scene and eye cameras, lenses, infrared light sources, IR pass filter, computer and power cables. The head mounted eye tracker is connected to a computer by USB port 2.0. An intermediary session uses a combining of wireless equipment.

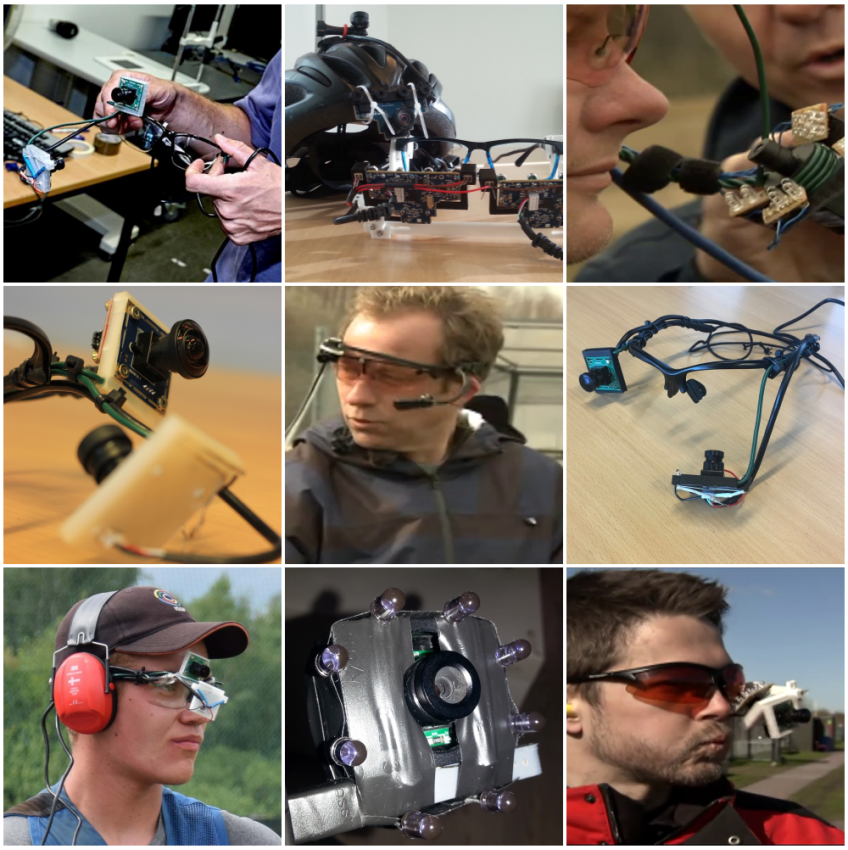


Figure 65: A set of head-mounted eye tracker prototypes built during this Ph.D. thesis.

Measures Measures and information collected during the shooting experiments evaluated the quality of cameras, infrared light sources, IR pass filters, m12 lenses with different focal lengths and apertures, components, among others. In the side of software, it was possible to assess the quality of captured images, maximum frame rate for capturing and recording videos, the efficiency of eye feature detectors, the accuracy and precision of gaze estimation, sources of noise, among others.

Procedure All shooting experiments were performed at the *Københavns Flugtskytte Klub* (KFK) where contains three Olympic fields of shooting practice with moving targets. Figure 66 shows the layout of a training field based on Olympic shooting rules. For

each experiment session, the athletes tried to hit the moving disk from the skeet range at stations 1, 4, 7 and 8. The session started with the personal calibration using nine targets. Then, the athlete performed three tasks, namely (1) hit five individual moving disks from the high tower; (2) hit five individual disks from the lower tower; and (i) hit two sequences of two disks released simultaneously from both towers, i.e. in total 14 targets for each trial. It was performed around ten training sessions of shooting experiments.

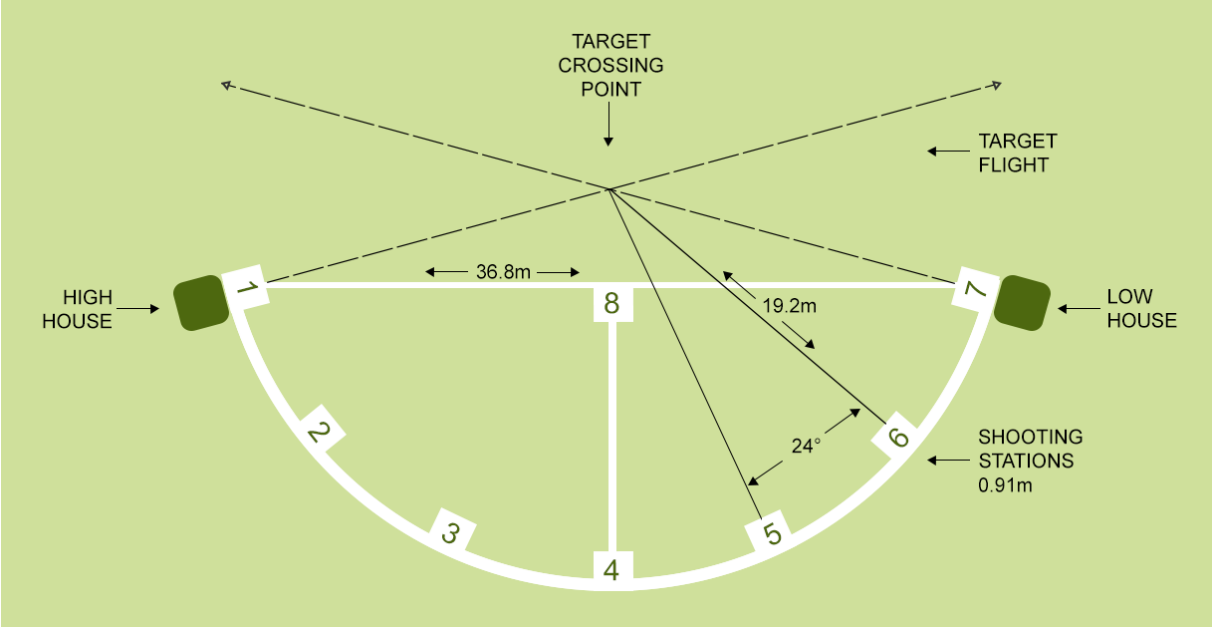


Figure 66: The layout of an Olympic field of shooting practice with moving targets.

7.2.2 Analysis and Results

After each daily training session, the collected gaze data were analyzed offline to investigate many aspects of both hardware and software (see Subsection 7.2.1). The prototypes built during this project were used to define the best off-the-shelf hardware to build a robust head-mounted eye tracker. For each session, the prototype was mounted on different supports, glasses frames, 3D printed mounts, helmets, and finally in safety glasses specific to shooting discipline, which presented the best ergonomy for shooting practices. In this sense, one of the main contributions of this Ph.D. thesis is to provide an entirely customizable eye tracker for different needs of each sports discipline.

The basic off-the-shelf hardware chosen to build eye trackers for sports experiments are (i) UVC cameras[1] that use m12 lenses, have board size of 38 × 38mm, are multi

[1]Available in https://goo.gl/VXq7ip

platform, work with USB2.0 high speed, and have maximum resolution of 1920 × 1080 Full HD in 30 frames per second using MJPEG codec; (ii) eye camera lenses[2] are 3.97mm (focal length), f/2.8 (aperture), 82 degrees (field of view), 16MP (resolution), with no distortion and have an infrared pass filter in a very narrow range around 850nm; and (iii) scene camera lenses[3] are 1.21mm, f/2.0, 220 degrees, 16MP, and they are able to capture a very wide field of view from the environment.

7.3 Kayak Experiment

This experiment had several aims in collaboration with the Department of Movement and Training Science of Natural Sports of the University of Leipzig and the German Kayak Organization. For the scope of this Ph.D. thesis, the main aims were (i) build an off-the-shelf head-mounted eye tracker adaptable to kayak discipline; and (ii) collect gaze data from a real eye tracking scenario of sport (*in-situ*) to be used in the improvement of eye tracking methods proposed in Chapter 6.

For the Leipzig's Associated Research Project[4], the main objective was using the collect gaze data to analyze the athlete's gaze behavior in canoe slalom during locomotion under competition constraints. The study investigated athletes with different experience levels in two whitewater kayak slalom disciplines, namely (1) canoe single [C-1], a sport which involves paddling a canoe with a single-bladed paddle; and (2) kayak single [K-1], with a double-bladed paddle.

The used head-mounted eye tracker prototype was built especially to the Leipzig's Associated Research Project. The research team had a first pilot test experiment using a commercial SMI head-mounted eye tracker. However, even *before* the athlete passed through the gate, it was not possible to see the two poles of the gate due to the narrow lens of the scene camera. For this reason, an adaptable head-mounted eye tracker was built to capture the entire whitewater slalom course.

7.3.1 Method

Participants A sample of 22 volunteer participants (13 males and 9 females) were recruited from the German Kayak National Team. Participants ranged from 18 to 31 years

[2]Available in https://goo.gl/qb2ZfX
[3]Available in https://goo.gl/4AuNiW
[4]See more information about this eye tracking study in the Experimental Test Plan available in Annex D.

old (23.68 ± 3.94). Eight athletes of canoe single (C-1) and fourteen athletes of kayak single (K-1). The athletes had different expertises in kayak practice, which six professionals, six experts, and ten intermediates. Their experience in kayak practice ranged from 8 to 22 years (15.00 ± 3.61). Three wore glasses and three wore contact lenses. Participants were free to blink, move the head, withdraw from testing at any stage.

Measures To the kayak experiment, the information collected during each session refers to eye and scene images, which we tried to find out solutions to improve the robustness of eye feature detection over critical outdoor illumination such as reflections, sun light, low contrast, among others. The Leipzig's Associated Research Project used this information to evaluate a qualitative approach to visual perception and a quantitative measurement of athletes' eye movements. Furthermore, a user evaluation was shared with this project to collect the user opinions related to (Q1) the eye tracking frame impaired the athlete's movements; (Q2) the situation of the study impaired the athlete's movements; and (Q3) the athlete's gaze movement matches with the real eye movements.

Apparatus The head-mounted eye tracker for kayak practice was mounted in the official canoe slalom helmet. The prototype had the following components (i) two synchronized Full-HD web cameras that allow capturing 60 frames per second with 1280×720 of resolution; (ii) one lens of 3.97mm (focal length) and f/2.8 (aperture) for the eye camera; (iii) one eye-fish lens of 1.21mm and f/2.0 for the scene camera; (iv) one IR narrow pass filter; and (v) some off-the-shelf hardware to mount electronic components in the helmet.

Pilot Tests The pilot tests consisted of four sessions and executed with one single kayak athlete. The objective was to evaluate the head-mounted eye tracker used in the kayak practices, speed up the capturing and recording video process, and to develop eye tracking methods robust to critical situations. For the pilot tests, the kayak athlete did not follow any formal experiment procedure, except the personal calibration. The pilot test was also used to define all steps of the formal follow experiment at the competition.

Procedure Two sessions were made at the actual test site in Augsburg, Germany. In the experiment, the athletes had to perform slalom courses in competition mode through a sequence of 7 gates. The gates have two poles of 1.2m width, numbered and colored as either red or green. Each color indicates the direction how the athlete must pass through the gate, namely (i) red (*upstream*), the athlete must rotate around one pole; and (ii) green (*downstream*), the athlete only must pass through the gate. For this experiment, only gates 1-7 were used, which 1-2 and 7 are red gates, and 3-6 are green.

Each session started by the personal calibration using twelve targets on a wide board

at 1.5 meters from the athlete. After that, the athlete went to the boat, and the laptop was placed inside the kayak with a water protection bag. Finally, the athlete performed the trials (each one around 45 seconds). The athlete repeated the experiment from 3 to 8 trials according to the number of failed trials (e.g. when the eye camera touched the gate pole). After the last trial, the personal calibration was verified by asking the participant to look at a moving target in nine positions.

At the end of the experiment, the athlete was interviewed to answer the evaluation questionnaire with three questions (Q1, Q2 and Q3) about the use of the head-mounted eye tracker. Posteriorly, the athletes answered about impressions and the results of processed scene videos. Figure 67 shows some images from the kayak experiment.

Figure 67: The kayak experiment with intermediate and expert athletes. (*top-left*) some tends have been placed along the river to reduce the influence of sun light during the experiment. (*top-middle*) the personal calibration with twelve targets. (*top-right*) the recalibration to check if the eye camera has moved during the experiment. (*bottom*) canoe and kayak athletes during the experiment.

7.3.2 Analysis and Results

Head Mounted Eye Tracker Through this experiment, we discovered a solution to build a stable head-mounted eye tracker for kayak practices, as shown in Figure 68. The eye and scene cameras were placed according to the camera location test (see Subsection 4.4.4), to reduce the influences of parallax error in X-axis. The list of used off-the-shelf hardware is available in Subsection 7.3.1.

Using the Proposed Eye Tracking Methods For the kayak experiment, it was necessary to develop eye feature detectors robust to the reflection in the water and direct

Figure 68: A head-mounted eye tracker prototype built with off-the-self hardware in a kayak helmet.

sun light, because the river acts as a mirror which reflects the sun light directly in the athletes' face. The eye tracking methods proposed in Chapter 6 played the main role in improving the gaze data collecting. Figure 69 shows the result of eye images processed with the eye tracking methods using priors.

Figure 69: Pupil detection of the same athlete and trial in different illumination conditions. Red X means pupils detected correctly and green X means pupils detected using eye information as priors.

7.3.3 User Satisfaction Evaluation

The user satisfaction evaluation had three questions (Q1, Q2 and Q3) related to the head-mounted eye tracker and the eye tracking session. Each athlete replied the questionnaire (available in Subsection 7.3.1) choose the best answer from 1 to 5, which (1) I do not agree; (2) I nearly do not agree; (3) I partly agree; (4) I nearly agree; and (5) I agree.

Figure 70 shows an overview of athletes' answers concerning the questionnaire. Q1 showed that the athletes answered that the head-mounted eye tracker impaired their natural movements (11/22 participants), and the second half said that the eye tracker partially impaired or did not impair. In general, Q2 showed that the athletes did not feel annoyed with the situation of the study (19/22 participants). Moreover, Q3 showed that most of the athletes wholly or partially agree on the gaze estimations matches with their real eye movements (17/22 participants).

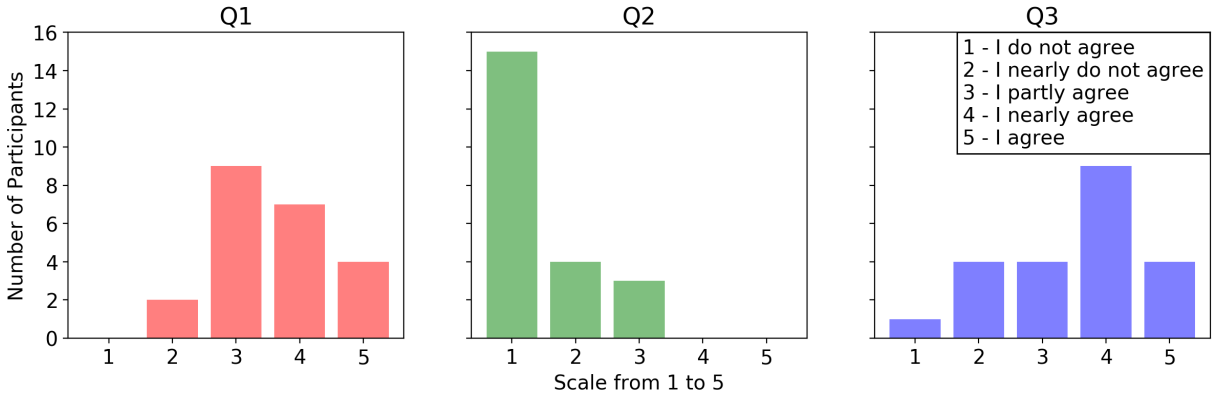


Figure 70: An overview of the user satisfaction evaluation from all 22 participants of the kayak experiment. They have answered the following satisfaction survey, Q1: Does the eye tracking frame impaired my movement? Q2: Does the situation of the study impaired my movement? and Q3: Does my gaze movements matches with my real eye movements?

The athletes that felt totally (18.2%) or partially (31.8%) impaired by the head-mounted eye tracker reported that (i) they need to take more distance to the gates or poles because their field of view was partially occluded or to avoid the eye camera touches the gate; (ii) they made an alternative trajectory and handling of the paddle because of helmet construction; (iii) paddling and movement were limited, and (iv) the frame distracted a little bit and the user touched it sometimes.

The athletes that did not feel annoyed (9.1%) or only a bit annoyed (40.9%) with the head-mounted eye tracker reported that (i) it was unusual to have a camera in their field of view, but they did not mind and (ii) they were more annoyed with the sun light changes and shadow under the tarp than with the eye tracker.

Some athletes from the group of women (9/22 participants) reported that the frame was too big, and the eye camera was in their field of view. It means, would be necessary to mount a female version of the head-mounted eye tracker to fit better in women's head. In the group of men (13/22 participants), some athletes reported the need to take more distance to the gates and the new strategies to pass through the gate.

Part IV

Conclusions and Future Work

Chapter

8

Conclusions

This Ph.D. thesis presents the use of *priors* to improve the accuracy and robustness of eye tracking system. The emphasis is on the application of eye tracking in sports, which is a challenging topic in the field of eye tracking due to the difficulty of performing eye tracking sessions in critical situations. The thesis embodies 4 main themes, namely (i) mathematical models to reduce the influence of noise in eye tracking systems; (ii) off-the-shelf hardware to build adaptable head-mounted eye trackers; (iii) eye tracking methods to explore the constraints revealed for specific sports settings; and (iv) prior information to improve the accuracy of gaze estimation methods. This chapter summarizes the contributions developed in this thesis and proposes future work.

8.1 Contributions of this Thesis

Using priors to improve eye tracking systems "*Using priors*" is a novel terminology presented in this Ph.D. thesis to the field of eye tracking, and it means to use available information known from the problem at hand as an alternative to improve different steps of eye tracking. Priors are the kernel of this Ph.D. thesis, and all eye tracking methods, models, and approaches developed in this Ph.D. thesis use prior knowledge from several data sources (e.g. orientation sensors, geometric relations of eye feature, depth, head angles). In general, eye tracking methods use eye feature purely to estimate the user's point of regard. This Ph.D. thesis goes beyond and uses any available information that could improve the robustness, accuracy, precision of eye tracking. The following topics present the kind of used priors and a statistical analysis of achieved improvements.

Building adaptable head-mounted eye trackers Most eye trackers suffer from geometric constraints as well as strong interference from the environment of the eye tracking session. Another factor to highlight is the low adaptability of commercial eye trackers

once it is not possible to replace electronic components according to constraints revealed for specific sports settings. For this reason, this thesis used head-mounted eye tracker prototypes built with off-the-shelf hardware that covers specific sports situations taking into account do not disturb the athlete's field of view during the daily training session.

Using the depth as priors to compensate the parallax error Section 4.2 presents a parallax compensation model that improves the gaze estimation accuracy in head-mounted eye trackers when the observed target moves in depth. The proposed compensation model uses the *distance* between the athlete and the observed target, and the *offset* of athlete's angle kappa as priors. The proposed model presented impressive results using both simulated and real gaze data. The evaluation using simulated gaze data showed that proposed model improves the accuracy by a factor of 10. Moreover, the evaluation using real gaze data demonstrated that the proposed model enhances the accuracy in 80.59%.

Using the 3D angles of head movements to compensate the head rotation Empirical observations of laboratory experiments revealed that head rotations influence the accuracy of head-mounted eye trackers. In general, the gaze error tends to the opposite direction concerning the head rotation. The magnitude and orientation of gaze error are user-dependent related to angle kappa, and it is necessary to learn about the error distribution for each subject. As head rotations are natural body movements, it is a need to create a mechanism to handle head rotation. An orientation sensor attached to the head-mounted eye tracker provides the three-dimensional head movements. The proposed head rotation compensation model uses the 3D angles from subject's head as priors to improve the gaze estimation accuracy by multiple linear regression. The evaluation in a real eye tracking scenario demonstrated that the proposed compensation model improves the gaze estimation accuracy in 86.41%.

Novel glint normalization approach This Ph.D. thesis presents a novel glint normalization approach based on homographies that uses a formal geometric correction scheme to make an eye tracking system more robust to head movements in remote eye tracking setups and more robust to eye camera slippage in head-mounted eye tracking setups. The proposed glint normalization approach is particularly focused on adapting the normalization according to the number of *detected* or *missing* corneal reflections. This approach requires labeling each detected glint to choose the correct mapping when the glint detection loses any true positive glint candidate. All geometric linear transformation degraded homographies depending on how many glints are available. The proposed approach introduces an alternative glint normalization that uses the N-closest glints from the pupil center (given

all true positive glints candidates) that reduces the influence of the non-linearly distribution of corneal reflections on the user's cornea surface. Experiments using simulated gaze data showed that the 3-closest and 2-closest glint normalization methods present a performance similar to traditional homography normalization [65] ($error \approx 0.4°$). This approach still needs to be evaluated in a real eye tracking scenario. The paper draft is available in Annex C.

Using eye feature as priors to compensate eye camera slippage Eye camera slippage is a common problem in mobile eye tracking setups, where the head and eye tracker moves and hence disrupting the personal calibration. For example in the shooting experiment, each shot fired the eye tracker vibrates and also changes the eye camera location. After the second shot, it was not possible to identify the actual athlete's point of regard in the scene image. The first trial of eye camera slippage compensation used corneal reflections to normalize the pupil center. However, glints were either weak due to outdoor lighting and disappeared due to large eyeball rotations. The proposed eye feature normalization approach uses the eye corners as reference points to normalize the pupil center. Experiments in a real eye tracking scenario showed that the proposed approach improves the gaze estimation accuracy in 74.94% compared to traditional gaze estimation without normalization. The eye corners detector developed in this Ph.D. thesis is still at an early stage of development and needs several improvements to stabilize the detected points along the processed video. This approach presented good results only in laboratory experiments (indoor), and it is exclusively for head-mounted eye tracking setup.

Using eyes as priors in a robust eye feature detector Eye tracking is challenging in critical situations which beyond geometric issues also influence eye feature extraction. The proposed eye feature detector uses the geometric relation of the human ocular system as priors to make eye tracking more robust to environmental noises. For example, the binocular eye tracking method (see Section 6.1) uses a geometric mapping between left and right eyes to validate and recognize both eyes in noised eye images. As the eye tracking data analysis evaluates offline gaze data, it is possible to have an overview of eyes behavior in a particular range at the eye tracking session time. The offline gaze data analysis is based on (i) N-neighborhood of the monocular eye tracking signal itself; and (ii) comparison of binocular eye tracking signals. This approach had a significant role in the analysis of gaze data from shooting and kayak experiments because sunlight made the pupil detection much more challenging. Lastly, the histogram analysis is used as priors to select automatically the threshold used for pupil and glints detectors. This approach works pretty well for the data set generated in this Ph.D. thesis and subsequent eye videos

recorded using the same eye tracker prototype. Before publishing the proposed approach, it is important to validate it using different datasets [33, 34, 128] with eye images from a larger number of participants performing different everyday tasks.

Eye tracking in sports experiments Experiments in different sports disciplines were performed to collect and analyze gaze data involving novice, intermediate and expert elite athletes during the daily training session of shooting, football, and kayak. Unfortunately, most of the eye tracking methods presented in this Ph.D. thesis were not tested in the sports experiments. In fact, the proposed eye tracking methods arose as a consequence of difficulties encountered in the sports experiments. All sports experiments were necessary to improve the developed eye tracking system and to build adaptable head-mounted eye trackers step-by-step. *Shooting experiments* assisted in defining cameras and lenses used in the final head-mounted eye tracker prototype, to build adaptable devices once the use of protection glasses is mandatory for shooting athletes in Denmark, and to define the best personal calibration to eye tracking in sports. *Football experiments* were useful to improve the video capturing and recording processes using multiple synchronized cameras and in the development of binocular eye tracking system. *Kayak experiments* were paramount to define the final head-mounted eye tracker prototype, capturing and recording videos in the maximum frame rate available in the selected cameras (60 Hz), and evaluation of binocular eye tracking methods presented in Section 6.1. The influence of parallax error, head rotation, and eye camera slippage was evaluated only in laboratory experiments.

8.2 Future Work

The research carried out in this Ph.D. thesis suggests some open problems and new aspects that could be further explored.

Using priors to improve eye tracking systems showed very compelling in sports experiments during this Ph.D. thesis. However, the best scenario would be to get priors from data sources available on the eye tracker itself. For example, extracting priors from captured scene images (i.e. object recognition/tracking) to estimate the distance between the user and the observed plane. This open problem will be the first research topic to be investigated after this Ph.D. thesis.

The proposed personal calibration is a bit complicated due to the inclusion of two extra calibration steps to compensate parallax error and head rotations. For uncalibrated head-mounted eye trackers, it is necessary to perform the personal calibration for the

same participant in each new eye tracking session. An approach to minimize and simplify the proposed personal calibration could be investigated.

The geometric relations of the human ocular system is the main parameter of a method to compensate and validate a binocular eye feature detector (see Subsection 6.1.1). However, the proposed method is restricted only to the pupil center. Further research in this direction could include a similar approach to detect missing glints based on a geometric transformation between corneal reflections available in both user's eyes.

Improvements in the eye tracker prototypes will be one of priority tasks of this research over the next few years. Nowadays, it is still necessary to use computers/laptops with good resources (e.g. SuperSpeed USB, multi core microprocessors, GPU, solid-state drives) to capture and record multiple synchronized videos in high resolution and good frame rate (i.e. more than 30 frames per second). The use of wireless setup to avoid occluding the subject's field of view during the eye tracking session could be further investigated.

Future steps in this research could also include using 360 scene cameras to capture images from the entire environment. This idea is useful in sports practices which the athlete remains in the same position throughout the sports practice. Thus, it would be possible to generate heat maps and gaze plots in different visualization perspectives which it would facilitate to compare the gaze behavior between a set of athletes.

References

[1] F. B. Narcizo and D. W. Hansen, "Depth compensation model for gaze estimation in sport analysis," in *Proceedings of the 2015 IEEE International Conference on Computer Vision Workshop*, ICCVW '15, (Washington, DC, USA), pp. 788–795, IEEE Computer Society, dec. 2015.

[2] Y. Wang, H. Zeng, and J. Liu, "Low-cost eye-tracking glasses with real-time head rotation compensation," in *Proceedings of the 10th International Conference on Sensing Technology*, ICST '16, (Washington, DC, USA), pp. 1–5, IEEE Computer Society, nov. 2016.

[3] T. B. Kinsman, *Semi-supervised pattern recognition and machine learning for eye-tracking*. PhD thesis, 1 Lomb Memorial Dr – NY 14623 – Rochester, USA, nov. 2015.

[4] J. N. Vickers and A. M. Williams, "Performing under pressure: the effects of physiological arousal, cognitive anxiety, and gaze control in biathlon," *Journal of Motor Behavior*, vol. 39, pp. 381–394, aug. 2007.

[5] J. Causer, P. S. Holmes, N. C. Smith, and A. M. Williams, "Anxiety, movement kinematics, and visual attention in elite-level performers," *Emotion*, vol. 11, pp. 595–602, mar. 2011.

[6] M. L. Mele and S. Federici, "Gaze and eye-tracking solutions for psychological research," *Cognitive Processing*, vol. 13, pp. 261–265, aug. 2012.

[7] D. Fegatelli, F. Giancamilli, L. Mallia, A. Chirico, and F. Lucidi, *The Use of Eye Tracking (ET) in Targeting Sports: A Review of the Studies on Quiet Eye (QE)*, pp. 715–730. Berlin, Heidelberg, Germany: Springer-Verlag, 2016.

[8] R. Paeglis, K. Bluss, A. Rudzitis, A. Spunde, T. Brice, and E. Nitiss, "Nir tracking assists sports medicine in junior basketball training," in *Proceedings of the International Society for Optical Engineering*, SPIE '11, (Bellingham, WA, USA), pp. 1–6, SPIE, jun. 2011.

[9] J. Afonso, J. Garganta, A. Mcrobert, A. M. Williams, and I. Mesquita, "The perceptual cognitive processes underpinning skilled performance in volleyball: Evidence

from eye-movements and verbal reports of thinking involving an in situ representative task," *Journal of Sports Science and Medicine*, vol. 11, pp. 339–345, jun. 2012.

[10] S. Hüttermann, D. Memmert, D. J. Simons, and O. Bock, "Fixation strategy influences the ability to focus attention on two spatially separate objects," *PLoS ONE*, vol. 8, pp. 1–8, jun. 2013.

[11] D. T. Bishop, N. Addington, and G. D'Innocenzo, "Using visual guidance to retrain an experienced golfer's gaze: A case study," *European Journal of Sport Science*, vol. 17, no. 2, pp. 160–167, 2017.

[12] D. L. Mann, W. Spratford, and B. Abernethy, "The head tracks and gaze predicts: how the world's best batters hit a ball," *PLoS ONE*, vol. 8, pp. 1–11, mar. 2013.

[13] R. Discombe and S. Cotterill, "Eye tracking in sport: A guide for new and aspiring researchers," *Sport and Exercise Psychology Review*, vol. 11, no. 2, pp. 49–58, 2015.

[14] A. Moran, A. Quinn, M. Campbell, B. Rooney, N. Brady, and C. Burke, "Using pupillometry to evaluate attentional effort in quiet eye: a preliminary investigation," *Sport, Exercise, and Performance Psychology*, vol. 5, pp. 365–376, nov. 2016.

[15] D. Panchuk and J. N. Vickers, "Using spatial occlusion to explore the control strategies used in rapid interceptive actions: predictive or prospective control?," *Journal of Sports Sciences*, vol. 27, pp. 1249–1260, oct. 2009.

[16] D. Panchuk, J. N. Vickers, and W. G. Hopkins, "Quiet eye predicts goaltender success in deflected ice hockey shots," *European Journal of Sport Science*, vol. 17, pp. 93–99, mar. 2017.

[17] D. W. Hansen and Q. Ji, "In the eye of the beholder: A survey of models for eyes and gaze," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 478–500, mar. 2010.

[18] J. Velez and J. D. Borah, "Visor and camera providing a parallax-free field-of-view image for a head-mounted eye movement measurement system," aug. 1989. U.S. Patent 4,852.988.

[19] D. Mardanbegi and D. W. Hansen, "Parallax error in the monocular head-mounted eye trackers," in *Proceedings of the 2012 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '12, (New York, NY, USA), pp. 689–694, ACM, sep. 2012.

[20] D. Mardanbegi, *Head-mounted gaze trackers for controlling the environment.* PhD thesis, IT University of Copenhagen, Rued Langgaards Vej 7 – DK-2300 – Copenhagen, Denmark, jul. 2013.

[21] D. Su, Y. F. Li, and C. Xiong, "Parallax error compensation for head-mounted gaze trackers based on binocular data," in *Proceedings of the 2016 IEEE International Conference on Real-time Computing and Robotics*, RCAR '16, (Washington, DC, USA), pp. 76–81, IEEE Computer Society, jun. 2016.

[22] D. Su and Y. F. Li, "Toward flexible calibration of head-mounted gaze trackers with parallax error compensation," in *Proceedings of the 2016 IEEE International Conference on Robotics and Biomimetics*, ROBIO '16, (Washington, DC, USA), pp. 491–496, IEEE Computer Society, dec. 2016.

[23] M. F. Land, "The coordination of rotations of the eyes, head and trunk in saccadic turns produced in natural situations," *Experimental Brain Research*, vol. 159, pp. 151–160, jun. 2004.

[24] B. Cesqui, R. van de Langenberg, F. Lacquaniti, and A. d'Avella, "A novel method for measuring gaze orientation in space in unrestrained head conditions," *Journal of Vision*, vol. 13, pp. 1–22, jul. 2013.

[25] T. Kocejko, A. Bujnowski, J. Ruminski, E. Bylinska, and J. Wtorek, "Head movement compensation algorithm in multi-display communication by gaze," in *Proceedings of the 7th International Conference on Human System Interactions*, HSI '17, (Washington, DC, USA), pp. 88–94, IEEE Computer Society, jun. 2014.

[26] S. M. Kolakowski and J. B. Pelz, "Compensating for eye tracker camera movement," in *Proceedings of the 2006 Symposium on Eye Tracking Research & Applications*, ETRA '06, (New York, NY, USA), pp. 79–85, ACM, mar. 2006.

[27] F. Li, S. M. Munn, and J. B. Pelz, "A model-based approach to video-based eye tracking," *Journal of Modern Optics*, vol. 55, no. 4-5, pp. 503–531, 2008.

[28] F. Karmali and M. Shelhamer, "Compensating for camera translation in video eye-movement recordings by tracking a representative landmark selected automatically by a genetic algorithm," *Journal of Neuroscience Methods*, vol. 176, pp. 157–165, jan. 2009.

[29] F. Li, *Optimizations and applications in head-mounted video-based eye tracking*. PhD thesis, 1 Lomb Memorial Dr – NY 14623 – Rochester, USA, 2011.

[30] D. W. Hansen and A. E. C. Pece, "Eye tracking in the wild," *Computer Vision and Image Understanding*, vol. 98, pp. 155–181, oct. 2005. Special Issue on Eye Detection and Tracking.

[31] W. J. Ryan, A. T. Duchowski, and S. T. Birchfield, "Limbus/pupil switching for wearable eye tracking under variable lighting conditions," in *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, ETRA '08, (New York, NY, USA), pp. 61–64, ACM, mar. 2008.

[32] J. B. Pelz, T. B. Kinsman, and K. M. Evans, "Analyzing complex gaze behavior in the natural world," in *Proceedings of the International Society for Optical Engineering*, SPIE '11, (Bellingham, WA, USA), pp. 78650Z–78650Z–11, SPIE, feb. 2011.

[33] L. Świrski, A. Bulling, and N. Dodgson, "Robust real-time pupil tracking in highly off-axis images," in *Proceedings of the 2012 Symposium on Eye Tracking Research & Applications*, ETRA '12, (New York, NY, USA), pp. 173–176, ACM, mar. 2012.

[34] W. Fuhl, T. C. Santini, T. Kübler, and E. Kasneci, "Else: ellipse selection for robust pupil detection in real-world environments," in *Proceedings of the 2016 Symposium on Eye Tracking Research & Applications*, ETRA '16, (New York, NY, USA), pp. 123–130, ACM, mar. 2016.

[35] M. M. Hayhoe, T. McKinney, K. Chajka, and J. B. Pelz, "Predictive eye movements in natural vision," *Experimental Brain Research*, vol. 217, pp. 125–136, mar. 2012.

[36] B. R. Pires, M. Hwangbo, M. Devyver, and T. Kanade, "Visible-spectrum gaze tracking for sports," in *Proceedings of the 1st IEEE International Workshop on Computer Vision in Sports*, CVsports '13, (Washington, DC, USA), pp. 1–6, IEEE Computer Society, jun. 2013.

[37] M. Dicks, C. Button, and K. Davids, "Examination of gaze behaviors under in situ and video simulation task constraints reveals differences in information pickup for perception and action," *Attention, Perception, & Psychophysics*, vol. 72, pp. 706–720, apr. 2010.

[38] J. Sigut and S.-A. Sidha, "Iris center corneal reflection method for gaze tracking using visible light," *IEEE Transactions on Biomedical Engineering*, vol. 58, pp. 411–419, feb. 2011.

[39] A. T. Duchowski, *Eye Tracking Methodology: Theory and Practice*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2nd edition ed., aug. 2007.

[40] J. Hnatow and A. Savakis, "An eye model for uncalibrated eye gaze estimation under variablehead pose," in *Proceedings of the International Society for Optical Engineering*, SPIE '07, (Bellingham, WA, USA), pp. 1–8, SPIE, apr. 2007.

[41] D. W. Hansen and R. I. Hammoud, "An improved likelihood model for eye tracking," *Computer Vision and Image Understanding*, vol. 106, pp. 220–230, may 2007.

[42] M. Reale, T. Hung, and L. Yin, "Pointing with the eyes: Gaze estimation using a static/active camera system and 3d iris disk model," in *Proceedings of the 2010 IEEE International Conference on Multimedia and Expo*, ICME '10, (Washington, DC, USA), pp. 280–285, IEEE Computer Society, jul. 2010.

[43] M. J. Reale, S. J. Canavan, L. Yin, K. Hu, and T. Hung, "A multi-gesture interaction system using a 3-d iris disk model for gaze estimation and an active appearance model for 3-d hand pointing," *IEEE Transactions on Multimedia*, vol. 13, pp. 474–486, jun. 2011.

[44] Z. Yuan and J. Kebin, "A local and scale integrated feature descriptor in eye-gaze tracking," in *Proceedings of the 4th International Congress on Image and Signal Processing*, CISP '11, (Washington, DC, USA), pp. 465–468, IEEE Computer Society, oct. 2011.

[45] S. Zhai, C. H. Morimoto, and S. Ihde, "Manual and gaze input cascaded (magic) pointing," in *Proceedings of the 1999 ACM SIGCHI Conference on Human Factors in Computing Systems*, SIGCHI '99, (New York, NY, USA), pp. 246–253, ACM, may 1999.

[46] F. da Silva Soares, "Mecanismos de interação ocular baseados em imagens voltados à inclusão digital de portadores de necessidades especiais," Master's thesis, Instituto Tecnológico de Aeronáutica, Praça Marechal Eduardo Gomes, 50 – 12.228-900 – São José dos Campos/SP, Brasil, nov. 2008.

[47] J. Chen, Y. Tong, W. D. Gray, and Q. Ji, "A robust 3d eye gaze tracking system using noise reduction," in *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, ETRA '08, (New York, NY, USA), pp. 189–196, ACM, mar. 2008.

[48] E. D. Guestrin and M. Eizenman, "Remote point-of-gaze estimation requiring a single-point calibration for applications with infants," in *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, ETRA '08, (New York, NY, USA), pp. 267–274, ACM, mar. 2008.

[49] E. D. Guestrin and M. Eizenman, "Remote point-of-gaze estimation with single-point personal calibration based on the pupil boundary and corneal reflections," in *Proceedings of the 24th IEEE Canadian Conference on Electrical and Computer Engineering*, CCECE '11, (Washington, DC, USA), pp. 971–976, IEEE Computer Society, may 2011.

[50] C. H. Morimoto, A. Amir, and M. Flickner, "Detecting eye position and gaze from a single camera and 2 light sources," in *Proceedings of the 16th International Conference on Pattern Recognition*, ICPR '02, (Washington, DC, USA), pp. 314–317, IEEE Computer Society, aug. 2002.

[51] C. Hennessey and P. D. Lawrence, "Improving the accuracy and reliability of remote system-calibration-free eye-gaze tracking," *IEEE Transactions on Biomedical Engineering*, vol. 56, pp. 1891–1900, jul. 2009.

[52] F. L. Coutinho and C. H. Morimoto, "A depth compensation method for cross-ratio based eye tracking," in *Proceedings of the 2010 Symposium on Eye Tracking Research & Applications*, ETRA '10, (New York, NY, USA), pp. 137–140, ACM, mar. 2010.

[53] K. Cen and M. Che, "Research on eye gaze estimation technique base on 3d model," in *Proceedings of the 2011 International Conference on Electronics, Communications and Control*, ICECC '11, (Washington, DC, USA), pp. 1623–1626, IEEE Computer Society, sep. 2011.

[54] F. L. Coutinho and C. H. Morimoto, "Augmenting the robustness of cross-ratio gaze tracking methods to head movement," in *Proceedings of the 2012 Symposium on Eye Tracking Research & Applications*, ETRA '12, (New York, NY, USA), pp. 59–66, ACM, mar. 2012.

[55] F. L. Coutinho and C. H. Morimoto, "Improving head movement tolerance of cross-ratio based eye trackers," *International Journal of Computer Vision*, pp. 1–23, 2012.

[56] Z. Zhu and Q. Ji, "Novel eye gaze tracking techniques under natural head movement," *IEEE Transactions on Biomedical Engineering*, vol. 54, pp. 2246–2260, dec. 2007.

[57] D. Torricelli, S. Conforto, M. Schmid, and T. D'Alessio, "A neural-based remote eye gaze tracker under natural head motion," *Computer Methods and Programs in Biomedicine*, vol. 92, pp. 66–78, oct. 2008.

[58] M. D. Phung, Q. V. Tran, K. Hara, H. Inagaki, and M. Abe, "Easy-setup eye movement recording system for human-computer interaction," in *Proceedings of the 2008 IEEE International Conference on Research, Innovation and Vision for the Future in Computing and Communication Technologies*, RIVF '08, (Washington, DC, USA), pp. 292–297, IEEE Computer Society, jul. 2008.

[59] A. Villanueva and R. Cabeza, "A novel gaze estimation system with one calibration point," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 38, pp. 1123–1138, aug. 2008.

[60] C. Zhang, J.-N. Chi, Z. Zhang, X. Gao, T. Hu, and Z. Wang, "Gaze estimation in a gaze tracking system," *Science China Information Sciences*, vol. 54, pp. 2295–2306, nov. 2011.

[61] C. H. Morimoto and M. R. M. Mimica, "Eye gaze tracking techniques for interactive applications," *Computer Vision and Image Understanding*, vol. 98, pp. 4–24, apr. 2005.

[62] F. B. Narcizo, "Sistemas de rastreamento ocular tolerantes aos movimentos da cabeça do usuário," relatório de projeto de tese, Universidade Federal de Campina Grande, Departamento de Sistemas e Computação, Universidade Federal de Campina Grande, BR, 2012.

[63] M. Barrett, H. Skovsgaard, and J. S. Agustin, "Performance evaluation of a low-cost gaze tracker for eye typing," in *Proceedings of the 5th Conference on Communication by Gaze Interaction*, COGAIN '09, (Lyngby, Denmark), pp. 13–17, The COGAIN Association, may 2009.

[64] J. San Agustin, H. Skovsgaard, E. Mollenbach, M. Barret, M. Tall, D. W. Hansen, and J. P. Hansen, "Evaluation of a low-cost open-source gaze tracker," in *Proceedings of the 2010 Symposium on Eye Tracking Research & Applications*, ETRA '10, (New York, NY, USA), pp. 77–80, ACM, mar. 2010.

[65] D. W. Hansen, J. S. Agustin, and A. Villanueva, "Homography normalization for robust gaze estimation in uncalibrated setups," in *Proceedings of the 2010 Symposium on Eye Tracking Research & Applications*, ETRA '10, (New York, NY, USA), pp. 13–20, ACM, mar. 2010.

[66] H.-C. Lu, G.-L. Fang, C. Wang, and Y.-W. Chen, "A novel method for gaze tracking by local pattern model and support vector regressor," *Signal Processing*, vol. 90, pp. 1290–1299, apr. 2010.

[67] H.-C. Lu, C. Wang, and Y.-W. Chen, "Gaze tracking by binocular vision and lbp features," in *Proceedings of the 19th International Conference on Pattern Recognition*, ICPR '08, (Washington, DC, USA), pp. 1–4, IEEE Computer Society, dec. 2008.

[68] C. Hennessey and P. D. Lawrence, "Noncontact binocular eye-gaze tracking for point-of-gaze estimation in three dimensions," *IEEE Transactions on Biomedical Engineering*, vol. 56, pp. 790–799, mar. 2009.

[69] Y. Huang, Z. Wang, and A. Ping, "Non-contact gaze tracking with head movement adaptation based on single camera," *World Academy of Science, Engineering and Technology*, vol. 59, pp. 395–398, nov. 2009.

[70] D. W. Hansen, L. Roholm, and I. G. Ferreiros, "Robust glint detection through homography normalization," in *Proceedings of the 2014 Symposium on Eye Tracking Research & Applications*, ETRA '14, (New York, NY, USA), ACM, mar. 2014.

[71] M. A. Just and P. A. Carpenter, "Eye fixations and cognitive processes," *Journal of Cognitive Psychology*, vol. 8, pp. 441–480, oct. 1976.

[72] M. Land and B. Tatler, *Looking and Acting: Vision and Eye Movements in Natural Behaviour*. OUP Oxford, 2009.

[73] F. L. Coutinho, "Um sistema de rastreamento de olhar tolerante a movimentação da face," Master's thesis, Universidade de São Paulo, Rua do Matão, 1010 – 05508-090 – Instituto de Matemática e Estatística, Cidade Universitária, São Paulo/SP, Brasil, may 2006.

[74] C.-N. Chan, S. Oe, and C.-S. Lin, "Active eye-tracking system by using quad ptz cameras," in *Proceedings of the 33rd Annual Conference of the IEEE Industrial Electronics Society*, IECON '07, (Washington, DC, USA), pp. 2389–2394, IEEE Computer Society, nov. 2007.

[75] J. J. Cerrolaza, A. Villanueva, M. Villanueva, and R. Cabeza, "Error characterization and compensation in eye tracking systems," in *Proceedings of the 2012 Symposium on Eye Tracking Research & Applications*, ETRA '12, (New York, NY, USA), pp. 205–208, ACM, mar. 2012.

[76] D. H. Yoo, J. H. Kim, B. R. Lee, and M. J. Chung, "Non-contact eye gaze tracking system by mapping of corneal reflections," in *Proceedings of the 5th IEEE International Conference on Automatic Face and Gesture Recognition*, FG '02, (Washington, DC, USA), pp. 94–99, IEEE Computer Society, may 2002.

[77] F. L. Coutinho and C. H. Morimoto, "Free head motion eye gaze tracking using a single camera and multiple light sources," in *Proceedings of the 19th Brazilian Symposium on Computer Graphics and Image Processing*, SIBGRAPI '06, (Washington, DC, USA), pp. 171–178, IEEE Computer Society, oct. 2006.

[78] D. H. Yoo and M. J. Chung, "A novel non-intrusive eye gaze estimation using cross-ratio under large head motion," *Computer Vision and Image Understanding*, vol. 98, pp. 25–51, apr. 2005.

[79] Y. Huang, Z. Wang, and X. Tu, "A real-time compensation strategy for non-contact gaze tracking under natural head movement," *Chinese Journal of Electronics*, vol. 19, pp. 446–450, jul. 2010.

[80] E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision*. Englewood Cliffs, NJ, USA: Prentice Hall, mar. 1998.

[81] I. D. Gluck, *Optics: The Nature and Applications of Light*. New York, NY, USA: Holt Rinehart & Winston, jan. 1964.

[82] D. T. Y. Mann, A. M. Williams, P. Ward, and C. M. Janelle, "Perceptual-cognitive expertise in sport: a meta-analysis," *Journal of Sport and Exercise Psychology*, vol. 29, pp. 457–478, aug. 2007.

[83] P. O'Donoghue, *An Introduction to Performance Analysis of Sport*. Abingdon-on-Thames, UK, England: Routledge, sep. 2014.

[84] J. M. D. Kremer, A. Moran, G. Walker, and C. Craig, *Key Concepts in Sport Psychology*. Thousand Oaks, CA, USA: SAGE Publications Ltd, 2012.

[85] P. O'Donoghue, *Research Methods Sports Performance Analysis*. Abingdon-on-Thames, UK, England: Routledge, feb. 2010.

[86] P. Passos, D. Araújo, and A. Volossovitch, *Performance Analysis in Team Sports*. Abingdon-on-Thames, UK, England: Routledge, dec. 2016.

[87] A. Roca, P. R. Ford, A. P. McRobert, and A. M. Williams, "Identifying the processes underpinning anticipation and decision-making in a dynamic time-constrained task," *Cognitive Processing*, vol. 12, pp. 301–310, aug. 2011.

[88] D. J. Hancock and D. M. Ste-Marie, "Gaze behaviors and decision making accuracy of higher- and lower-level ice hockey referees," *Psychology of Sport and Exercise*, vol. 14, pp. 66–71, jan. 2013.

[89] P. Vansteenkiste, R. Vaeyens, L. Zeuwts, R. Philippaerts, and M. Lenoir, "Cue usage in volleyball: a time course comparison of elite, intermediate and novice female players," *Biology of Sport*, vol. 4, no. 31, pp. 295–302, 2014.

[90] D. Alder, P. R. Ford, J. Causer, and A. M. Williams, "The coupling between gaze behavior and opponent kinematics during anticipation of badminton shots," *Human Movement Science*, vol. 37, pp. 167–179, oct. 2014.

[91] A. L. Yarbus, *Eye Movements and Vision*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., sep. 1967.

[92] J. P. Pluijms, R. C. nal Bruland, M. J. M. Hoozemans, M. W. V. Beek, K. Böcker, and G. J. P. Savelsbergh, "Quantifying external focus of attention in sailing by means of action sport cameras," *Journal of Sports Sciences*, vol. 34, pp. 1588–1595, nov. 2016.

[93] R. Schmidt and T. Lee, *Motor Control and Learning: A Behavioral Emphasis*. Champaign, IL, USA: Human Kinetics, 5th edition ed., mar. 2011.

[94] M. W. Eysenck, N. Derakshan, R. Santos, and M. G. Calvo, "Anxiety and cognitive performance: attentional control theory," *Emotion*, vol. 7, pp. 336–353, May 2007.

[95] N. F. Fogt and A. B. Zimmerman, "A method to monitor eye and head tracking movements in college baseball players," *Optometry and Vision Science*, vol. 2, pp. 200–211, feb. 2014.

[96] A. Iwatsuki, T. Hirayama, and K. Mase, "Analysis of soccer coach's eye gaze behavior," in *Proceedings of the 2013 Asian Conference on Pattern Recognition*, IAPR '13, (Washington, DC, USA), pp. 793–797, IEEE Computer Society, nov. 2013.

[97] M. J. Campbell and A. P. Moran, "There is more to green reading than meets the eye! exploring the gaze behaviours of expert golfers on a virtual golf putting task," *Cognitive Processing*, vol. 15, pp. 363–372, aug. 2014.

[98] J. Afonso, J. Garganta, A. McRobert, M. Williams, and I. Mesquita, "Visual search behaviours and verbal reports during film-based and in situ representative tasks in volleyball," *European Journal of Sport Science*, vol. 14, pp. 177–184, oct. 2014.

[99] T. Heinen, K. Velentzas, and P. M. Vinken, "Functional relationships between gaze behavior and movement kinematics when performing high bar dismounts – an exploratory study," *Human Movement*, vol. 13, pp. 218–224, oct. 2012.

[100] G. Wood and M. R. Wilson, "A moving goalkeeper distracts penalty takers and impairs shooting accuracy," *Journal of Sports Sciences*, vol. 28, pp. 937–946, may 2010.

[101] L. J. Moore, S. J. Vine, P. Freeman, and M. R. Wilson, "Quiet eye training promotes challenge appraisals and aids performance under elevated anxiety," *International Journal of Sport and Exercise Psychology*, vol. 11, pp. 169–183, apr. 2013.

[102] J. Spitz, K. Put, J. Wagemans, A. M. Williams, and W. F. Helsen, "Visual search behaviors of association football referees during assessment of foul play situations," *Cognitive Research: Principles and Implications*, vol. 1, pp. 1–11, oct. 2016.

[103] O. Neumann and A. F. Sanders, *Handbook of Perception and Action: Attention.* Cambridge, MA, USA: Academic Press, may 1996.

[104] J. N. Vickers, "Origins and current issues in quiet eye research," *Current Issues in Sport Science*, vol. 1, pp. 1–11, 2016.

[105] L. J. Moore, S. J. Vine, M. R. Wilson, and P. Freeman, "The effect of challenge and threat states on performance: An examination of potential mechanisms," *Psychophysiology*, vol. 49, pp. 1417–1425, aug. 2012.

[106] M. R. Wilson and R. C. Pearcy, "Visuomotor control of straight and breaking golf putts," *Perceptual and Motor Skills*, vol. 109, pp. 555–562, oct. 2009.

[107] L. Fischer, R. Rienhoff, J. Tirp, J. Baker, B. Strauss, and J. Schorer, "Retention of quiet eye in older skilled basketball players," *Journal of Motor Behavior*, vol. 47, pp. 407–414, aug. 2015.

[108] M. R. Wilson, G. Wood, and S. J. Vine, "Anxiety, attentional control, and performance impairment in penalty kicks.," *Journal of Sport and Exercise Psychology*, vol. 6, pp. 761–75, dec. 2009.

[109] S. Bernet, P. Sturm, C. Cudel, and M. Basset, "Study on the interest of hybrid fundamental matrix for head mounted eye tracker modeling," in *Proceedings of the 2011 British Machine Vision Conference*, BMVC '11, (South Road, Durham, UK), pp. 15.1–15.10, BMVA Press, aug. 2011.

[110] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, UK, England: Cambridge University Press, 2nd edition ed., mar. 2004.

[111] H. C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," *Nature*, vol. 293, pp. 133–135, sep. 1981.

[112] S. J. D. Prince, *Computer Vision: Models, Learning, and Inference*. Cambridge, UK, England: Cambridge University Press, jul. 2012.

[113] M. Böhme, M. Dorr, M. Graw, T. Martinetz, and E. Barth, "A software framework for simulating eye trackers," in *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, ETRA '08, (New York, NY, USA), pp. 251–258, ACM, mar. 2008.

[114] C. H. Morimoto, D. Koons, A. Amir, and M. Flickner, "Frame-rate pupil detector and gaze tracker," in *Proceedings of the 7th IEEE International Conference on Computer Vision*, ICCV '99, (Washington, DC, USA), IEEE Computer Society, sep. 1999.

[115] Q.-T. Luong and O. D. Faugeras, "The fundamental matrix: theory, algorithms, and stability analysis," *International Journal of Computer Vision*, vol. 17, pp. 43–75, jan. 1996.

[116] J. F. Collins and L. K. Blackwell, "Effects of eye dominance and retinal distance on binocular rivalry," *Perceptual and Motor Skills*, vol. 39, pp. 747–754, oct. 1974.

[117] S. Alfayad, M. E. Asswad, A. Abdellatif, F. B. Ouezdou, A. Blanchard, N. Beaussé, and P. Gaussier, "HYDROïD humanoid robot head with perception and emotion capabilities: modeling, design, and experimental results," *Frontiers in Robotics and AI*, vol. 3, pp. 1–15, apr. 2016.

[118] P. L. Kaufman and A. Alm, *Adler's Physiology of the Eye*. Philadelphia, PA, USA: Mosby, 10nd edition ed., jul. 2003.

[119] E. Hering, *The Theory of Binocular Vision*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 1977.

[120] C. W. Oyster, *The Human Eye: Structure and Function*. Sunderland, MA, USA: Sinauer Associates Inc., U.S., feb. 2006.

[121] R. Biedert, J. Hees, A. Dengel, and G. Buscher, "A robust realtime reading-skimming classifier," in *Proceedings of the 2012 Symposium on Eye Tracking Research & Applications*, ETRA '12, (New York, NY, USA), pp. 123–130, ACM, mar. 2012.

[122] A. Auinger, P. Brandtner, P. Großdeßner, and A. Holzinger, "Search engine optimization meets e-business – a theory-based evaluation: findability and usability as key success factors," in *Proceedings of the 2012 International Conference on Data Communication Networking, e-Business and Optical Communication Systems*, DCNET, ICE-B and OPTICS '12, (Lisboa, Portugal), pp. 237–250, SciTePress, jul. 2012.

[123] J. C. de Carvalho Melo, F. B. Narcizo, D. W. Hansen, C. Brabrand, and A. Wasowski, "Variability through the eyes of the programmer," in *Proceedings of the 25th International Conference on Program Comprehension*, ICPC '17, (Washington, DC, USA), pp. 34–44, IEEE Computer Society, may 2017.

[124] X. Wei, P. Lucey, S. Morgan, P. Carr, M. Reid, and S. Sridharan, "Predicting serves in tennis using style priors," in *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, SIGKDD '15, (New York, NY, USA), pp. 2207–2215, ACM, aug. 2015.

[125] K. Holmqvist, M. Nystrom, R. Andersson, R. Dewhurst, H. Jarodzka, and J. van de Weijer, *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford, UK: Oxford University Press, mar. 2015.

[126] D. W. Hansen, M. Nielsen, J. P. Hansen, A. S. Johansen, and M. B. Stegmann, "Tracking eyes using shape and appearance," in *Proceedings of the 2002 Conference on Machine Vision Applications*, MVA '02, (Nara, Japan), pp. 201–204, Nara-Ken New Public Hall, dec. 2002.

[127] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proceedings of the of 2014 IEEE Computer Vision and Pattern Recognition*, CVPR '14, (Washington, DC, USA), pp. 1867–1874, IEEE Computer Society, jun. 2014.

[128] W. Fuhl, T. Kübler, K. Sippel, W. Rosenstiel, and E. Kasneci, "Excuse: Robust pupil detection in real-world scenarios," in *Proceedings of the 16th International Conference Computer Analysis of Images and Patterns*, CAIP '15, (Berlin, Heidelberg, Germany), pp. 39–51, Springer-Verlag, sep. 2015.

Annex

A

Full paper published in CVsports 2015 at ICCV 2015

Depth Compensation Model for Gaze Estimation in Sport Analysis

Fabricio Batista Narcizo and Dan Witzner Hansen
IT University of Copenhagen
Rued Langgaards Vej 7, 2300 København S., Danmark
{fabn,witzner}@itu.dk

Abstract

A depth compensation model is presented as a novel approach to reduce the effects of parallax error for head-mounted eye trackers. The method can reduce the parallax error when the distance between the user and the target is prior known. The model is geometrically presented and its performance is tested in a totally controlled environment with aim to check the influences of eye tracker parameters and ocular biometric parameters on its behavior. We also present a gaze estimation method based on epipolar geometry for binocular eye tracking setups. The depth compensation model has shown very promising to the field of eye tracking. It can reduce 10 times less the influence of parallax error in multiple depth planes.

1. Introduction

Eye tracking is an expressive tool for analyzing human interest and intend [13]. Its potential for detailed and objective performance analysis in sports has been shown in various experiments [1, 12, 17] and the collected information from eye tracking can potentially help athletes to become more effective in their daily training.

Head mounted eye tracking is the most obvious type of eye tracker to use for analysis of athletes during daily training [8] and with current technological developments it could be implemented without significantly disturbing the athletes. However, head mounted eye tracking analysis are typically challenged by the parallax error, which happens as a consequence of the spatial offset between the eye and the camera observing the scene [15]. This causes significant gaze estimation errors (in the scene view) when the apparent objects are located at different depths than during calibration. In addition to the inherent inaccuracy of gaze estimation (0.5 degrees or more), the parallax error will effectively make it hard, if not impossible, to analyze the eye tracking data reliably. Despite of this, many research results are based on manual inspection of video data with overlaid gaze data and the depth compensated point of regard is done based on human estimation when the objects move in space. The parallax error is a practical problem for gaze-analysis in sports where the target (say ball, stone or person) constantly moves in depth relative to the athlete. So far there is no commercial eye tracker that can account for the parallax error and therefore research results are often based on human inspection and estimates on the location the point of regard (PoR) when the target moves in space [13].

In this paper we will present a method that uses the depth as a prior to compensate for the parallax error. Having the object depth can be made feasible via visual tracking or through other sensors. Even without an accurate estimate the method can be used to discern between which of two objects the person is most likely to look at. In Section 2 we describe related work and in Section 3 we describe the parallax error in more detail. The parallax error compensation model is described in Section 4.

The proposed depth compensation model is shown to consistently improve the accuracy level of gaze estimation process when the target is viewed on both calibration plane (Section 5.2) and different depth planes (Section 5.3). Through this paper we intend to show that it is possible to estimate the athletes' gaze actively in given sports situations and thus overcome some of the problems relate to gaze estimation in depth using head-mounted eye trackers. An overview of eye and gaze tracking models is reviewed by Hansen and Ji [8].

2. Related Work

Eye tracking has been used for sports analysis but mostly using head mounted eye tracking [1, 10, 12, 14, 17, 18]. Most eye tracking results are of psychological nature but gaze estimation data collected during daily training where it can be used for analyzing the athletes' performance, such as what happens when the athlete perform specific actions (e.g., shoot, catch, throw)? Are there ocular differences between novice and experts [12, 14, 17]? Which strategies can be used to improve the novice athletes' performance based on knowledge of eye movements patterns collected from training activities of expert athletes [1]? For example,

Paeglis *et al*. [17] presented a study for analyzing eye movements from elite junior basketball players during shots practice training. After only a year of training, their free shot rates significantly improved as a consequence of the use of eye tracking in their training. They concluded among other important results that if compared directly with expert basketball players, novice basketball players need more time for quality decisions before making their shots but during free shots, expert players spend more time to do this action than novice players.

Hüttermann *et al*. [12] presented a study for verifying the ability to devote attention simultaneously to multiple visual objects into a athlete's field of view and important concept in team sports, for example. Hüttermann *et al*. [12] showed that athletes present better attention performance when focusing attention simultaneously on two stimuli (i.e., when the athlete fixates between two stimuli). This could mean, in the future, it will be possible to improve training athletes to better focus through of subsequent training studies for improving fixation strategies of a determinate athlete.

Eye trackers for sports analysis are based on the same fundamental eye and gaze tracking models as described in the overview [8]. The parallax error in head-mounted eye trackers (HMET) can be minimized through hardware. Velez and Borah [20] presented an eye tracker that uses a hot-mirror glass in front of the user's eyes for controlling the distances and angular relationships of eye camera, scene camera and user's eyes. The hot-mirror is positioned in the eye's optic axis with aim to reflect images from the eyes and environment toward their respective cameras. This setup removes the parallax error and ensures a wide-angle scene viewing over multiple depth planes. Not all HMET have a physical structure that allows hot-mirror glasses to be used. Mardanbegi and Hansen [15] proposed a study to identify the main sources of parallax error in head mounted eye trackers. They analyzed the influence of scene camera positions, the calibration and fixation distances on the parallax error. They showed that the angle *kappa* (the difference between visual and optical axes) does not have a significant effect on the parallax error [15].

3. Parallax Error in Gaze Estimation

We are going to explain the parallax error based on a HMET. The most of current HMET cannot estimate high accuracy gaze due to the parallax error. Parallax error is a geometrical problem due to the projection center of HMET scene camera and the user's eyeball center are not co-axial. Since the scene camera cannot be placed at the same axis of the user's eye, it is necessary to find out a solution for compensating the parallax error for general HMET.

HMET usually have two (monocular) or three (binocular) cameras attached in their physical structure. For example, a binocular HMET has two eye cameras located slightly close to each eye and one scene camera on the head. The scene camera is used for capturing images from the user's field of view. However, the scene camera is usually not co-axial with any user's eye. In this case, the gaze estimation includes the parallax error on it. Figure 1 shows an example of the parallax error when a person uses an HMET like that.

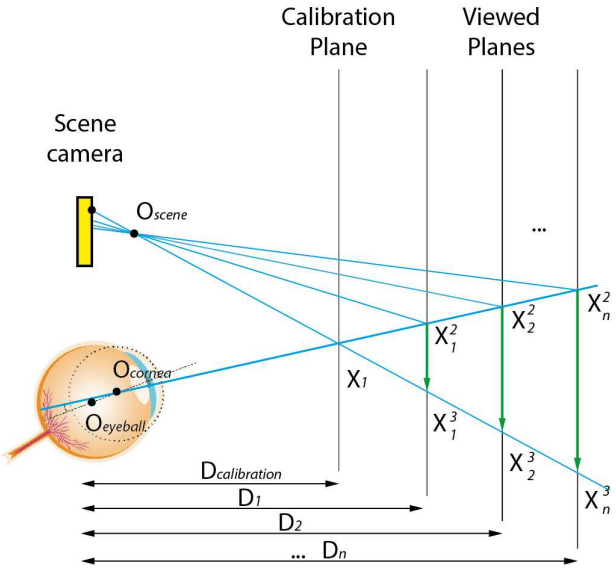


Figure 1. Geometry of parallax error in a HMET. The head mounted eye tracking system is calibrated on the calibration plane. Targets on calibration plane can be estimated with high accuracy. However, when the user looks at to a target in the same position X_1 but on a different plane D_i, the gaze will be estimated on position X_i^3 instead of the position X_i^2. The green arrows represent the parallax error corresponding to the vector $||X_i^2 X_i^3||$ on the viewed plane and a vector $||x^1 x_i^2||$ on the scene camera plane.

The first parameter to be analyzed is the user's eye. According to *Gullstrand-Le Grand Eye Model*, the simplified model for representing the human visual system is formed by two spheres with distinct sizes for representing the eyeball and the cornea surface [4, 5, 11]. The center of rotation of these spheres is around a fixed point $O_{eyeball}$ and there is a small angular difference between the optical and visual axes, which is user dependent and they intersect in the point O_{cornea}. The second parameter is the scene camera, in which is not co-axial with the user's eye. The scene camera is represented as a pinhole camera with a vertical image plane. The last parameter to be analyzed are the planes viewed by the user during the eye tracking session.

The eye tracking system is calibrated in a given distance $D_{calibration}$ from the user to the calibration plane. All points on the calibration plane can be estimated with high accuracy level. However, what happen when the user fixates his/her gaze so far away from the calibration plane (at

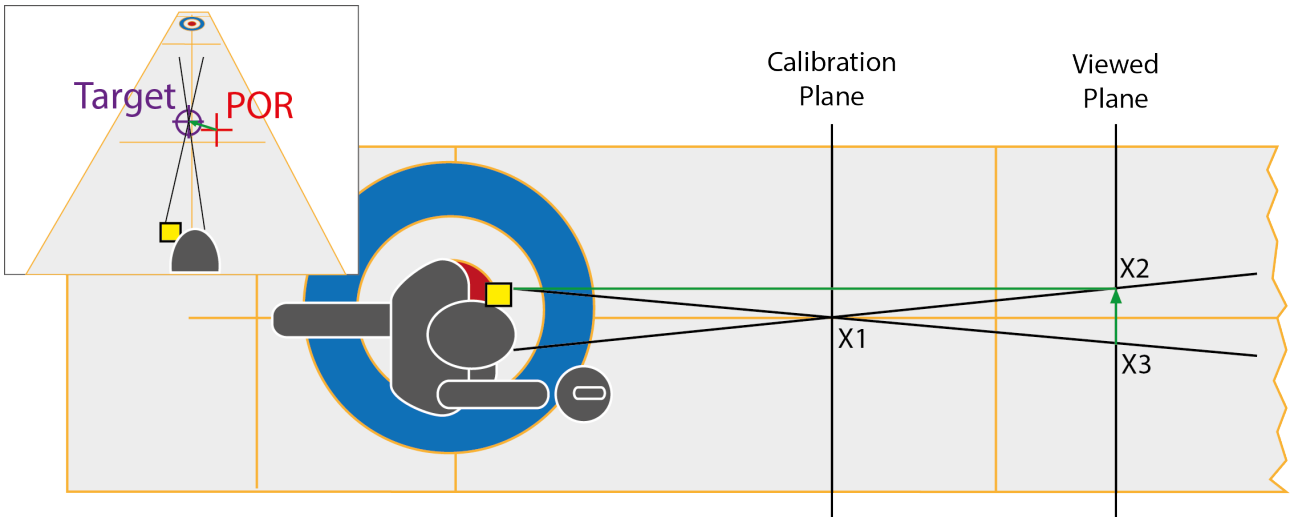


Figure 2. Example of parallax error during a curling daily training session. The system is calibrated on the calibration plane and there is no influence of parallax error on that. On the other hand, when the stone goes to position X_2 on the viewed plane, the gaze will be estimated on position X_3. When the distance between the curling athlete and the viewed stone is prior known, the proposed depth compensation model is able to correct the estimated PoR for the correct position in the scene image, as shown in the small upper picture.

the distance D_i)? The user's visual axis intersects the calibration plane exactly at the point X^1 and the multiple depth planes at the point X_i^2 $(1 \leq i \leq n)$. When the user looks directly at the point X^1, the gaze is estimated correctly as the point x^1 on the image plane. On the other hand, when the user looks at to any depth plane at the point X_i^2, the gaze is going to be estimated as the same point x^1 instead of the point x_i^2 on the image plane. According to Mardanbegi and Hansen (2012), the parallax error is defined as a vector $||x^1 x_i^2||$ on the image plane, which is corresponding to the vector $||X_i^2 X_i^3||$ on the viewed plane [15]. Figure 2 shows a practical example of using an eye tracking system for a curling daily training. The parallax error appears in different planes on the curling sheet when the athlete look at to the stone (target) far way from the calibration plane.

4. Depth Compensation Model

The phenomenon of parallax is related by the geometry of the HMET, in which it can be described by *epipolar geometry* in a stereo vision system [2, 15]. In this case, the epipolar geometry is expressed by the point p_{eye} on the eye plane and the point p_{scene} on scene camera plane that must lie on a line called *epipolar line*. As shown in Figure 3, if the point P in the athlete's field of view moves along the line formed by the optical center of the scene camera O_{scene} and the point p_{scene}, its projection on the scene plane will not change but the projection on the eye plane will change. This movement traces out the epipolar line D_{eye} [2].

When the athlete focuses on objects at different planes (see Figure 1), the PoR projection will move along an epipolar line in the image plane. Based on epipolar geometry, all epipolar lines intersect at a common point called *epipole*. In our context, the epipole is placed in the optical center $O_{eyeball}$ into eye plane. Each epipolar line can be estimated through an algebraic representation called *fundamental matrix* (F). F can be estimated given at least seven point correspondences in both image and scene camera planes. These correspondences represent the geometric information about the intrinsic and extrinsic parameters of the cameras.

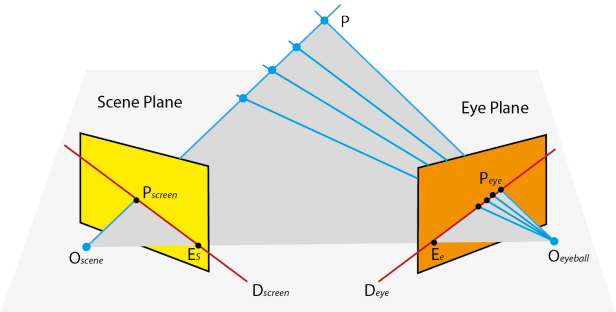


Figure 3. Geometry of parallax error in a HMET.

The fundamental matrix F encapsulates the intrinsic camera geometry and it is independent of scene structure. Given F as a 3×3 matrix, it is possible to calculate the corresponding epipolar line D_{screen} for every point p_{eye} in the eye image plane by Equation 1:

$$D_{screen} = F \times p_{eye}. \quad (1)$$

If any point P is imaged as p_{eye} in the eye camera and

p_{screen} in the scene camera, then the following relation is equivalent to the corresponding epipolar line and constrains the matching of points through Equation 2:

$$p_{screen}^{T} \times F \times p_{eye} = 0, \tag{2}$$

for all corresponding points $p_{screen} \leftrightarrow p_{eye}$.

In a binocular HMET, each pair of eye camera and scene camera forms a structure similar to a stereo vision system. We uses a binocular eye tracking approach to define two epipolar lines in the scene camera plane. As the PoR moves along each epipolar line, the intersection between these epipolar lines will be very close to the real athlete's gaze. The biggest problem is when the target moves so far away from the calibration plane, because the parallax error will drastically reduce the precision of the gaze estimation process. Our depth compensation model is based on pure translation motion, i.e. a planar motion case where there is no rotation [9, 19]. According to Hartley and Zisserman [9], the "*pure*" definition means that there is no change in the internal parameters of the cameras.

After estimating the gaze, it is necessary to correct the parallax error based on the information about the depth distance between the viewed target and the calibration plane. As the cameras are stationary in the HMET and we consider that only the targets undergoes a translation $-t$. In this case, the three-dimensional points move on straight lines parallel in the direction of t. One may assume that the calibration plane and viewed plane are respectively $P_{calibration} = K[I|0]$ and $P_{viewed} = K'[I|t]$. Equation 3 is used for calculating the fundamental matrix when there is no rotation ($R = I$) and both camera matrices are the same ($K = K'$):

$$F = [e']_{\times} K' K^{-1} = [e']_{\times} \tag{3}$$

in which the notation $[e']_{\times}$ is a rank 2 skew-symmetric 3×3 matrix. If the target plane translation is parallel to the calibration plane z-axis, then $e' = (0, 0, 1)^T$. In this case, the fundamental matrix can be represented by the Equation 4:

$$F = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{4}$$

If a point x in the calibration plane is normalized as $x = (x, y, 1)^T$, then from $x = P_{calibration} X = K[I|0]X$, the space point's coordinates are $(X, Y, Z)^T = ZK^{-1}x$, where Z is the depth of the point X from the viewed plane along the principal axis of the calibration plane. It then follows from $x' = P'X = K[I|t]X$, Equation 5 corrects the estimated gaze point x to the real gaze point x' viewed by the athlete without parallax error:

$$x' = x + Kt/Z \tag{5}$$

in which depends on the magnitude value of the translation t and the inverse depth Z [9].

5. Assessment on Simulated Data

Simulated eye tracking data were used for assessing the proposed gaze estimation approach (using epipolar geometry) and depth compensation model (using pure translation) in a totally controlled environment. We have used a MATLAB eye tracker simulator in which it is possible to control the eye tracker parameters and the ocular biometric parameters [3]. Therefore, it was possible to evaluate the noise effects of each parameter in the gaze estimation process. The evaluation process was divided according to when the user visualizes targets on the calibration plane (Subsection 5.2) and on the multiple depth planes (Subsection 5.3).

The assessment on the calibration plane has evaluated the accuracy of the proposed gaze estimation approach based on epipolar geometry, and the assessment on the multiple depth planes has evaluated the parallax error rectification of the proposed depth compensation model based on pure translation. We have evaluated the following aspects during our assessment process: (1) refractive index of aqueous humor [α]; (2) number of calibration targets [N]; (3) horizontal [γ] and vertical [β] angle offset between optical and visual axes [a.k.a. *angle kappa*]; (4) the influence of noise in the eye features detection process [$Pc + \lambda$]; and (5) depth movements along to the calibration plane z-axis.

5.1. Setup

The simulated eye tracker device was setup as a binocular HMET. In this case, it had two eye cameras (one for each user's eye) located slightly close to the calibration plane and one scene camera (to get images from the user's field of view) slightly close to the user's head. The calibration plane was adjusted to 55 cm distance from the user. During each test, it was estimated the gaze error from 4,096 targets distributed in a 64×64 matrix over the viewed plane. For all simulated tests, we have used two eye models with different angle kappa offsets [7], namely: $E_0[\beta = \gamma = 0°]$ (a physically infeasible setup only to avoid some eye specific biases) and $E_1[\beta = 1.5°, \gamma = 4.5°]$ (a more realistic ocular biometric setting). For standard, it has used the minimum number of calibration targets necessary to create the fundamental matrix ($N = 8$) during the user calibration process.

5.2. Tests on the Calibration Plane

Refractive Index of Aqueous Humor The first test evaluated the influence of refraction index in the gaze estimation process. According to Hansen and Ji [8], the refractive index of aqueous humor has a constant value around 1.336. It can add some noise or have some directly influence to the gaze estimation process. Table 1 presents the influence of the refractive index of aqueous humor when it is included and when it is not. We concluded that there is no influence of the refractive index in the gaze estimation process. On

the other hand, changes in the angle kappa offset present a notable difference for a similar test.

Model	Refraction	Maximum Error	Mean Error
E_0	No	$0.0000231°$	$0.0000037°$
E_0	Yes	$0.0000231°$	$0.0000037°$
E_1	No	$0.0791455°$	$0.0088679°$
E_1	Yes	$0.0791455°$	$0.0088679°$

Table 1. The influence of refractive index of aqueous humors [1.336] to the gaze estimation using E_0 [$\beta = \gamma = 0°$] and E_1 [$\beta = 1.5°, \gamma = 4.5°$] eye models.

Number of Calibration Targets The second test evaluated how the number of calibration targets ($8 \leq N \leq 25$) influences the gaze estimation process. The user calibration process is performed in the beginning of an eye tracking session. The user needs to look at N targets on the calibration plane for creating a mapping used by the gaze estimation process. Figure 4 shows the accuracy of the gaze estimation as a function of the number of calibration targets. For both eye models the minimum number of calibration targets ($N = 8$) has achieved a good accuracy. In opposite to the classical gaze estimation methods, in a more realistic ocular biometric setting this approach does not improve its accuracy when the number of the calibration targets increase during the user calibration process (see the blue graphic).

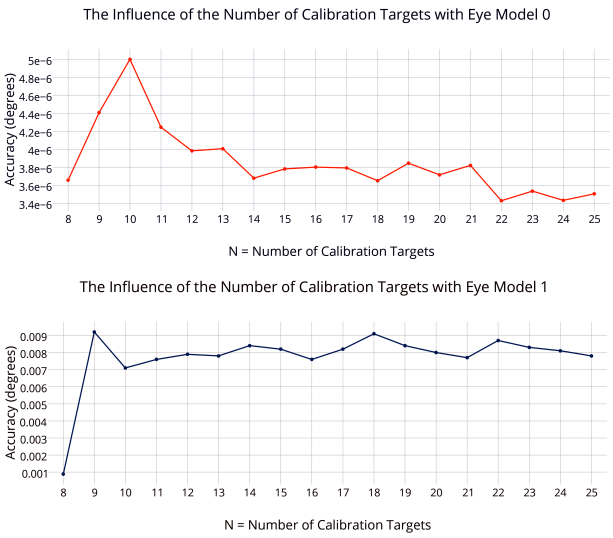


Figure 4. The influence of the number of calibration targets N to the gaze estimation using eye model (up) $E_0[\beta = \gamma = 0°]$ and (down) eye model E_1 [$\beta = 1.5°, \gamma = 4.5°$].

Angle Kappa Offset The third test showed that there is a huge difference among the tests performed with different angle kappa offsets. Figure 5 shows the influence of different angle kappa offsets within a range of angular horizontal offsets ($-4.5° \leq \gamma \leq 4.5°$) and a fixed angular vertical offset ($\beta = 0°$). We conclude that the gaze estimation based on epipolar geometry does not model the angle kappa with high accuracy. The accuracy linearly decreases according to angle kappa, i.e. the bigger the angular difference among visual and optical axes the lower will be the accuracy.

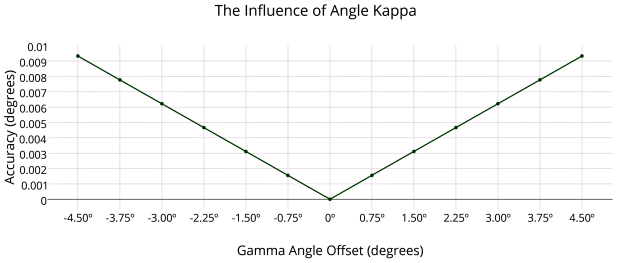


Figure 5. The influence of the angle kappa offset to the gaze estimation. Angle kappa has two angles offsets, i.e. horizontal (γ) and vertical (β). We observed the influence of angle kappa with $-4.5° \leq \gamma \leq 4.5°$ and $\beta = 0°$.

Noise During the aforementioned tests, the gaze estimation approach based on the intersection of multiple epipolar lines showed very promising ($error < 0.01°$). However, in a real application will this approach achieve the same accuracy degree? With aim to answer this question, we have performed a fourth test add a controlled noise in the pupil center coordinate before calculate the epipolar line, i.e. $line_{left} = F_{left} \times (Pc_{left} + \lambda)$ and $line_{right} = F_{right} \times (Pc_{right} + \lambda)$. Figure 6 shows a two-dimensional view of the noise tests with different values to horizontal coordinates and a fixed vertical coordinate $Pc = (x + \lambda, y)$.

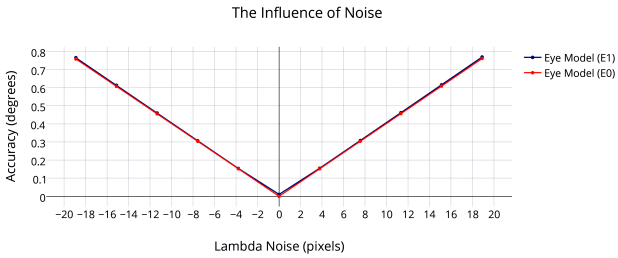


Figure 6. Two-dimensional view of the influence of noise added to the pupil center coordinate to the gaze estimation process using eye model (up) E_0 [$\beta = \gamma = 0°$] and (down) eye model E_1 [$\beta = 1.5°, \gamma = 4.5°$]. The noise ($\lambda$) was added to $P_{center} = (x, y)$ in the following range $-18.90 \leq \lambda \leq 18.90$ pixels.

We concluded that this gaze estimation approach only calculate the PoR with high accuracy level when there is no noise in the eye features detection process. In a real eye tracking application, this approach is going present the same accuracy degree as the classical gaze estimation meth-

ods (e.g. homography, cross-ratio, polynomial, among others [5, 7, 16, 23]). Figure 7 shows a three-dimensional views of the influence of noise in the gaze estimation process. The noise was added to each (x, y) coordinate of the pupil center, in the following range $-18.90 \leq \lambda \leq 18.90$ pixels (± 5 mm) on the calibration plane.

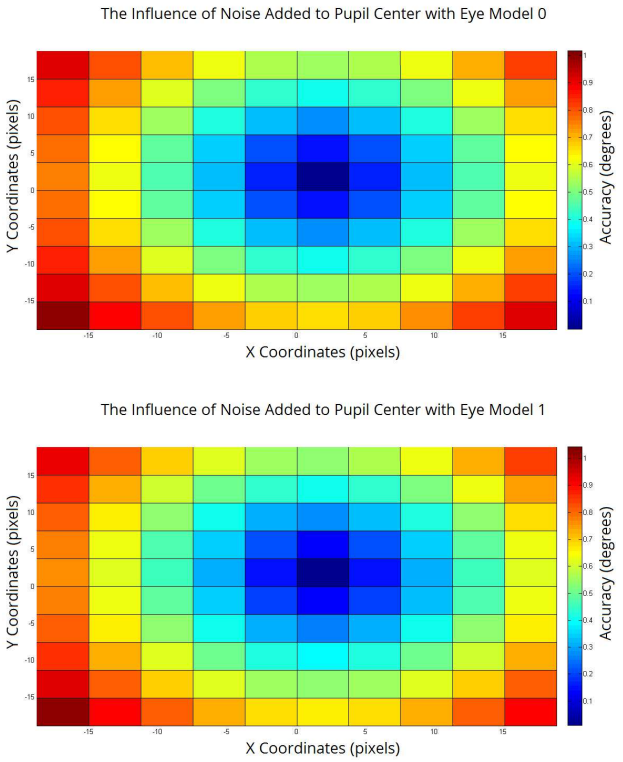


Figure 7. Three-dimensional view of the influence of noise added to the pupil center coordinate to the gaze estimation process using eye model (up) E_0 $[\beta = \gamma = 0°]$ and (down) eye model E_1 $[\beta = 1.5°, \gamma = 4.5°]$. The noise (λ) was added to $P_{center} = (x, y)$ in the following range $-18.90 \leq \lambda \leq 18.90$ pixels.

There is a huge difference only when the noise is $\lambda = 0$. In this case, E_0 presents an accuracy degree around $0.0106921°$ and E_1 is around $0.0000039°$. For others noise level, the accuracy degree is basically the same for E_0 and E_1, i.e. the difference mean is $\pm 0.01°$.

5.3. Tests on the Depth Planes

The last simulated test was performed with aim to evaluate the depth compensation model proposed in this paper. During this test, the HMET hardware components and the user are still while the targets moves along to the calibration plane z-axis. The calibration plane is 55 cm far away from the user and the viewed plane moves in a range 35-105 cm from the user (step of 10 cm). Figure 8 and 9 show the influence of parallax error to the gaze estimation. For each 10 cm far away from the calibration plane, the parallax error adds a gaze error around $\pm 0.23°$.

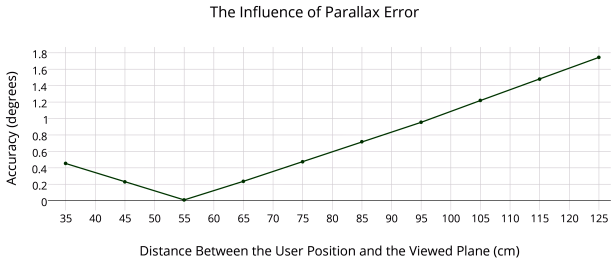


Figure 8. The influence of parallax error to the gaze estimation. The viewed plane was moved to 10 different distances far away from the user position, i.e. in a range 35-125 cm. The accuracy level decrease because the parallax error.

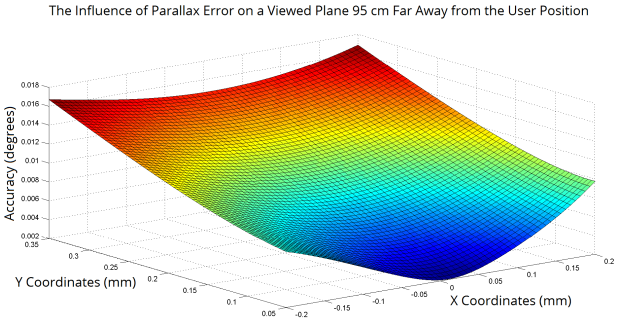


Figure 9. The influence of parallax error to the gaze estimation. After the user calibration process, the viewed plane was moved to 95 cm far away from the user position. The gaze estimation presented an accuracy level around $0.95°$.

Figure 10 and 11 show the influence of proposed depth compensation model to the gaze estimation process. The depth compensation model was able to correct the parallax error of simulated eye tracking data. For each 10 cm far away from the calibration plane, the parallax error adds only a gaze error around $\pm 0.02°$ (i.e. 10 times less). At this point, we concluded that this depth compensation model is very promising to the field of eye tracking.

6. Conclusions

This paper has presented a novel depth compensation model used for correcting the parallax error in HMET. The proposed model is robust to large depth planes when the distance between the user and the target is prior known. The distance is used for compensating the parallax error using the pure translation approach. This paper has also described a gaze estimation method based on epipolar geometry. This method has presented high accuracy degree with simulated date. However, it has shown very sensitive to intrinsic and

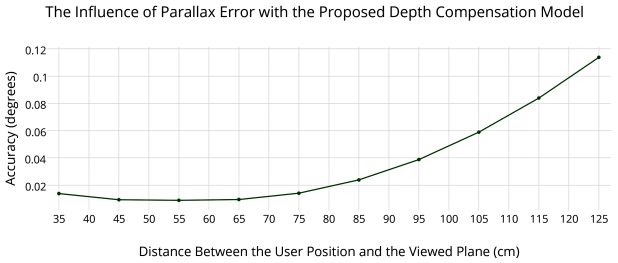


Figure 10. The influence of depth compensation model to the gaze estimation. The viewed plane was moved to 10 different distances far away from the user position, i.e. in a range 35-125 cm. The gaze estimation presented a good accuracy level despite the parallax error.

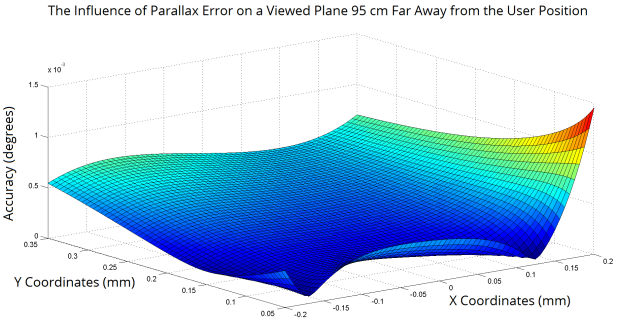


Figure 11. The influence of depth compensation model to the gaze estimation. Using the depth compensation model, the gaze estimation achieved an accuracy level around $0.04°$ when the viewed plane was moved to 95 cm far away from the user position.

extrinsic noise and it accuracy is similar to others classical gaze estimation methods (e.g. homography, cross-ratio, polynomial, among others).

The proposed model was developed to be used in a bigger project with elite sport athletes (shooting and curling). The main expected contributions by this research project is to develop flexible eye tracking models that can be used for elite sport athletes in their daily training. Eye tracking has been used for sports and has already shown some promise. However, eye trackers used for sports analysis are general purpose, expensive and not adapted to be used actively in sports situations.

Eye tracking can used for collecting information about the pattern of ocular activities of experts athletes and let other novices athletes observe their eye movements. The use of eye tracking in sport can go further, e.g. to auxiliary the hawk-eye technology for evaluating information that has raised doubts during a match [22], to activate resources of a vehicle cockpit through fixations [21], to find the better alternative to view multiple targets during an action of attack or defenses [12] and to identify external points of distraction presents during an eye tracking session [6]. Eye tracking data and tools will allow the athletes and trainers to get much deeper insight into thoughts and strategies used by the athletes, and adapt the training correspondingly thus improving their performance in stressful and time critical situations. While the focus of this project is on sports training, it is evident that progress made within this project on eye tracking and supporting tools for sports activities could have a direct impact on other areas that use eye tracking.

Acknowledgments

The authors would like to thank CNPq – Brazilian National Council for Scientific and Technological Development – for the financial support to the first author in his PhD research project.

References

[1] J. Afonso, J. Garganta, A. Mcrobert, A. M. Williams, and I. Mesquita. The perceptual cognitive processes underpinning skilled performance in volleyball: Evidence from eye-movements and verbal reports of thinking involving an in situ representative task. *Journal of Sports Science and Medicine*, 11(1):339–345, jun. 2012.

[2] S. Bernet, P. Sturm, C. Cudel, and M. Basset. Study on the interest of hybrid fundamental matrix for head mounted eye tracker modeling. In *Proceedings of the 2011 British Machine Vision Conference*, BMVC '11, pages 15.1–15.10, South Road, Durham, UK, aug. 2011. BMVA Press.

[3] M. Böhme, M. Dorr, M. Graw, T. Martinetz, and E. Barth. A software framework for simulating eye trackers. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, ETRA '08, pages 251–258, New York, NY, USA, mar. 2008. ACM.

[4] K. Cen and M. Che. Research on eye gaze estimation technique base on 3d model. In *Proceedings of the 2011 International Conference on Electronics, Communications and Control*, ICECC '11, pages 1623–1626, Washington, DC, USA, sep. 2011. IEEE Computer Society.

[5] F. L. Coutinho and C. H. Morimoto. Improving head movement tolerance of cross-ratio based eye trackers. *International Journal of Computer Vision*, pages 1–23, 2012.

[6] T. M. Gable, B. N. Walker, H. R. Moses, and R. D. Chitloor. Advanced auditory cues on mobile phones help keep drivers' eyes on the road. In *Proceedings of the* 5^{th} *International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '13, pages 66–73, New York, NY, USA, oct. 2013. ACM.

[7] D. W. Hansen, J. S. Agustin, and A. Villanueva. Homography normalization for robust gaze estimation in uncalibrated setups. In *Proceedings of the 2010 Symposium on Eye Tracking Research & Applications*, ETRA '10, pages 13–20, New York, NY, USA, mar. 2010. ACM.

[8] D. W. Hansen and Q. Ji. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3):478–500, mar. 2010.

[9] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, UK, England, 2^{nd} edition edition, mar. 2004.

[10] M. M. Hayhoe, T. McKinney, K. Chajka, and J. B. Pelz. Predictive eye movements in natural vision. *Experimental Brain Research*, 217(1):125–136, mar. 2012.

[11] C. Hennessey and P. D. Lawrence. Improving the accuracy and reliability of remote system-calibration-free eye-gaze tracking. *IEEE Transactions on Biomedical Engineering*, 56(7):1891–1900, jul. 2009.

[12] S. Hüttermann, D. Memmert, D. J. Simons, and O. Bock. Fixation strategy influences the ability to focus attention on two spatially separate objects. *PLoS ONE*, 8(6):1–8, jun. 2013.

[13] M. Land and B. Tatler. *Looking and Acting: Vision and Eye Movements in Natural Behaviour*. OUP Oxford, 2009.

[14] D. L. Mann, W. Spratford, and B. Abernethy. The head tracks and gaze predicts: how the world's best batters hit a ball. *PLoS ONE*, 8(3):1–11, mar. 2013.

[15] D. Mardanbegi and D. W. Hansen. Parallax error in the monocular head-mounted eye trackers. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 689–694, New York, NY, USA, sep. 2012. ACM.

[16] C. H. Morimoto, D. Koons, A. Amir, and M. Flickner. Frame-rate pupil detector and gaze tracker. In *Proceedings of the 7^th IEEE International Conference on Computer Vision*, ICCV '99, Washington, DC, USA, sep. 1999. IEEE Computer Society.

[17] R. Paeglis, K. Bluss, A. Rudzitis, A. Spunde, T. Brice, and E. Nitiss. Nir tracking assists sports medicine in junior basketball training. In *Proceedings of the International Society for Optical Engineering*, SPIE '11, pages 1–6, Bellingham, WA, USA, jun. 2011. SPIE.

[18] B. R. Pires, M. Hwangbo, M. Devyver, and T. Kanade. Visible-spectrum gaze tracking for sports. In *Proceedings of the 1^st IEEE International Workshop on Computer Vision in Sports*, CVsports '13, pages 1–6, Washington, DC, USA, jun. 2013. IEEE Computer Society.

[19] S. J. D. Prince. *Computer Vision: Models, Learning, and Inference*. Cambridge University Press, Cambridge, UK, England, jul. 2012.

[20] J. Velez and J. D. Borah. Visor and camera providing a parallax-free field-of-view image for a head-mounted eye movement measurement system, aug. 1989. U.S. Patent 4,852.988.

[21] H. Vrzakova and R. Bednarik. Hard lessons learned: mobile eye-tracking in cockpits. In *Proceedings of the 4^th Workshop on Eye Gaze in Intelligent Human Machine Interaction*, Gaze-In '12, pages 7:1–7:6, New York, NY, USA, oct. 2012. ACM.

[22] X. Wei, P. Lucey, S. Morgan, P. Carr, M. Reid, and S. Sridharan. Predicting serves in tennis using style priors. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, SIGKDD '15, pages 2207–2215, New York, NY, USA, aug. 2015. ACM.

[23] D. H. Yoo and M. J. Chung. A novel non-intrusive eye gaze estimation using cross-ratio under large head motion. *Computer Vision and Image Understanding*, 98(1):25–51, apr. 2005.

Annex

B

Abstract accepted in ECEM 2017

Abstract of Contribution 611

Wednesday, 23/Aug/2017 11:50am - 12:10pm
ID: 611 / Wed1ThemSessR3: 5
Talk preferred
Topics: Low cost eye tracking, 2D/3D eye tracking, Mobile eye tracking
Keywords: parallax error, head rotations, head-mounted eye tracker, sports analysis

Using Priors to Compensate Geometrical Problems in Head-Mounted Eye Trackers

Fabricio Batista Narcizo, Zaheer Ahmed, Dan Witzner Hansen
IT University of Copenhagen, Denmark

The use of additional information (a.k.a. priors) to help the eye tracking process is presented as an alternative to compensate classical geometrical problems in head-mounted eye trackers. Priors can be obtained from several distinct sources, such as: sensors to collect information related to distance, location, luminance, movement, speed; information extracted directly from the scene camera; calibration of video capture devices and other components of the eye tracker; information collected from a totally controlled environment; among others. Thus, priors are used to improve the robustness of eye tracking in real applications, for example, (1) If the distance between the subject and the viewed target is known, it is possible to estimate subject's current point of regard even when target moves in depth and suffers influence of parallax error; and (2) if the tridimensional angular rotation is known, it is possible to compensate the error induced by the head rotations using linear regression. Experiments with simulated eye tracking data and in real scenarios of elite sports have been showing that the use of priors to support the eye tracking systems help produce more accurate and precise gaze estimation specially for uncalibrated head-mounted setups.

Annex

C

Draft paper to be submitted

Multiple Geometrical Transformations for Robust Glint Normalization

Fabricio Batista Narcizo*
IT University of Copenhagen

Dan Witzner Hansen†
IT University of Copenhagen

Abstract

A novel glint normalization approach based on multiple geometrical transformations for uncalibrated remote eye trackers is presented. This approach selects the best geometrical transformation according to the number of detected corneal reflections. This makes glint normalization more robust to noise in a real eye tracking application. We demonstrate geometrically and empirically that this approach is robust to head pose changes and it maintains high-accuracy gaze estimation even when only one glint is detected.

CR Categories: I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Hierarchy and geometric transformations;

Keywords: glint normalization, homographies, remote eye tracker

1 Introduction

This paper addresses glint normalization in the presence of missing glints. Glint normalization is a method that uses a formal geometric correction scheme for making an eye tracking system more robust to head movements for a remote eye tracking setup [Hansen et al. 2014]. The paper is particularly focused on adapting glint normalization according to the number of detected glints.

Corneal reflections (a.k.a glints) are commonly used in scientific researches in the field [Coutinho and Morimoto 2012; Hansen et al. 2010] and they can support invariant head pose changes to traditional gaze estimation methods (e.g. PCCR, homography, cross-ratio). They can be generated through *active* or *passive* illumination. Active illumination is usually controlled by the eye tracker and it is very common to use infrared light sources (IR). On the other hand, passive illumination uses the environment natural illumination or reflections of environment objects. It is more common to use active illumination for generating glints because it is easier to control these illumination sources (e.g. localization, intensity).

While head-mounted eye trackers do benefit from corneal reflections [Li et al. 2008], it is mostly remote eye trackers that need to handle head pose changes. When the eyeball moves around its center of rotation, the corneal reflections from IR sources remain still. However, glint detection often fails for large angles of eye rotation when the corneal reflections occur close to the limbus (the boundary between the iris and the sclera) or on the sclera itself (the white part of the eye). Glint normalization is typically challenged by noise and eye pose changes, which reduces the number of true positive corneal reflections. This induces gaze estimation errors when the glints positions are not detected with high accuracy, the glint intensity is so weak, the glint is reflected on the sclera, some required glint is not detected, uncontrolled light environment conditions or there are several false positives. In summary, glint normalization can be cornerstones in making remote eye tracking systems more robust to natural head movements [Hansen et al. 2014].

In Section 2 we describe related work and in Section 3 we introduce the proposed glint normalization approach based on multiple geometrical transformations. This approach is shown to consistently maintain the robustness of gaze estimation when a different number of glints are detected on both head still (see Subsection 4.2) and head movements (see Subsection 4.3). Through this paper we intend to show that it is possible to maintain the robustness of an eye tracking system even with large head movements and thus overcome some of the influence of noise relate to glint normalization using remote eye trackers. An overview of eye tracking methods, technologies and applications is given in Hansen and Ji [2010].

2 Previous Work

Several eye tracking researches were presented in order to make the glint normalization more robust to external critical situations. Hua et al. [2006] introduce a glint detection approach that handles addresses the missing glint problem over large angles of eye rotation. Infrared light sources placed in a strategic position form a fixed geometric pattern of glints and this geometrical condition is predictable when the eye rotates arbitrarily. Based on a prior known geometric pattern, it is possible to calculate the intersection point (a.k.a. virtual glint) with less than 1 pixel error even with some missing glints. This method works with a fixed number of glints (i.e. four glints) and it can estimate the virtual glint when one or two of the four glints were missed. Li et al. [2008] propose a similar glint detection approach for predicting missing glints. This method uses a fixed geometric pattern of glints formed by nine infrared light sources (3×3 matrix) and it is able to predict missing glints when the size of the pattern remains relatively fixed. It can predict missing glints when at least one glint is detected on the cornea region. Hennessey and Lawrence [2009] present an enhanced Pupil Center Corneal Reflection (PCCR) method that uses a scaling correction for estimating the PCCR vector. It uses Euclidean distance between the pupil center and the centroid of the glints geometric pattern and it can improve the gaze estimation 2.8 times more (for near depth head movements) compare to traditional PCCR (using one glint). This method compensates for translation, distortion, false positives and missing glints. However, the method requires at least two valid glints for a robust glint normalization. Hansen et al. [2014] demonstrate a high accuracy glint detection method that detect the corneal reflections even in the presence of several false positives. This methods uses homography normalization and it is based on geometric properties of corneal reflections that learn about the unknown configuration of IR for detecting only the glints generated by the system. In this case, the glint detection is robust up to around 10 false positives reflections without using other constraints than the fixed geometric pattern.

3 Proposed Glint Normalization Method

This section presents a novel glint normalization approach based on homography and degraded homographies for uncalibrated remote eye tracker setups. The proposed approach is towards an eye tracker

*e-mail: fabn@itu.dk
†e-mail: witzner@itu.dk

setup with one remote camera and four infrared light sources, as shown in Figure 1. The simplified geometric model of the human eye contains the point p_c as the center of the elliptical pupil, the point C as the center of the cornea sphere and f_c as the cornea focal point. Degraded homographies are used to estimate the relation between points from image space Π_i and points from screen space Π_s. Thus, given corresponding points ($x_i \Leftrightarrow x'_i$) on two planes (Π and Π'), the geometrical transformation is a matrix $M_{3\times3}$ such that for any point $x'_i = M \times x_i$.

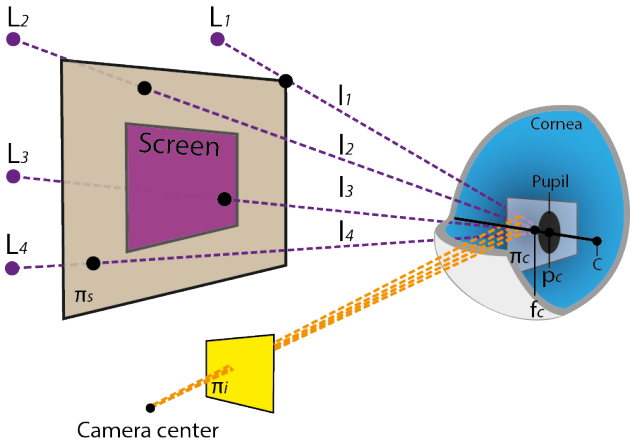


Figure 1: *Geometric setup of the remote camera, human eye intrinsic parameters, infrared light sources and corneal reflections.*

Our proposed model uses among one and four corneal reflections ($g^c_1 ... g^c_4$) from the infrared light sources on the cornea. These reflections are formed in a plane called corneal-reflection plane Π_c. Differently from cross-ratio based models [Yoo and Chung 2005; Coutinho and Morimoto 2012], the proposed model do not need to place the infrared light sources exactly on the corners of the screen. We assume that the infrared light sources are simply placed on the screen plane Π_s, and the relationship between the infrared light sources (L_i) and their respective reflections (g^c_j) can be related via a geometrical transformation M^s_c. The remote camera is placed in front of the eyes and it captures a geometric pattern of glints ($g^i_1 ... g^i_4$) on the image plane Π_i. The relationship between the corneal reflections (g^c_j) on the corneal-reflection plane Π_c and the geometric pattern of glints (g^i_j) on the image plane Π_i can be related via a geometrical transformation M^i_c.

Our gaze estimation method is described by a direct mapping from the image plane Π_i to the screen plane Π_s using $M^s_i = M^c_i \circ M^s_c$. Furthermore, it uses a normalized space Π_n between the image space (Π_i) and screen space (Π_s) for normalized the pupil center (p_c) with the proposed glint normalization approach based on multiple geometrical transformation, as shown in Figure 2. Thus, the pupil center and glints are normalized from *image space* to *normalized space* via a dynamic glint normalization transformation H^n_i, and finally, the normalized pupil center (p^n_c) is used for estimating the gaze on the screen via a fixed homography F^s_n.

Eye tracking techniques based on active illumination present some problems in relation to glint detection, e.g. bad illumination distribution, glints collision, insufficient bright, large angles of eye rotation, among others. For traditional homography normalization methods, it is necessary to detect all four available glints for a robust gaze estimation. However, sometimes the glint detection does not detect all true positive glints. The proposed approach works with multiple geometrical transformations that adapt the glint normalization for normalizing points from the space Π_i to the space Π_n using $1 \leq j \leq 4$ glints. Its basic requirement is to label each detected glint for helping to choose the correct mapping when the glint detection loses any true positive. All geometrical transformation matrices are created during the user calibration process and they are dynamically selected by the glint normalization according to the number of detected glints.

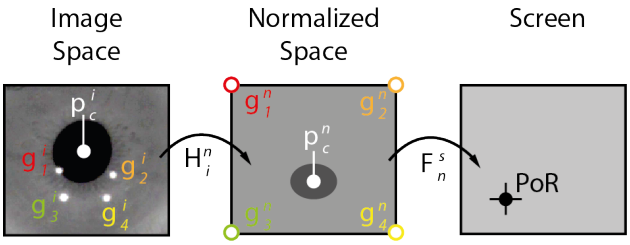


Figure 2: *(left) Pupil (gray ellipse) and glints (crosses) are captured in the image space and (middle) the pupil is normalized using the detected glints by a geometrical transformation. (right) From the normalized space the pupil is mapped to the Point of Regard.*

The proposed approach uses different geometrical transformations with distinct Degrees of Freedom (DoF). In general, linear transformations can be written as an invertible 3×3 matrix in homogeneous coordinates with real entries, as defined in Equation 1:

$$M_p = \left[\begin{array}{cc|c} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ \hline v_1 & v_2 & r \end{array}\right] = \begin{bmatrix} A & t \\ v^T & r \end{bmatrix}, \qquad (1)$$

in which, A denotes a 2×2 non-singular matrix, t denotes a translation vector, v denotes perspective components and r denotes homogeneous scaling factor [Hartley and Zisserman 2004].

Our proposed glint normalization is based on four geometrical transformations, namely: (1) *Translation*, a transformation in which all segments on the transformed plane are identical to their respective segments on the original plane but in a different position. It maintains the distance and angle of all transformed points and it can be computed from only one detected glint; (2) *Similarity*, a transformation composed by an isometry transformation (translation and rotation in the same measure) and a isotropic scaling factor. It preserve the shape (i.e. no reflection) and it can be computed from two detected glints correspondences; (3) *Affine*, a non-singular linear transformation of inhomogeneous coordinates followed by a translation. It does not preserve the length ratios and angles between lines because its non-isotropic scaling and it can be computed from three detected glints correspondences; and (4) *Homography*, a general non-singular linear transformation of homogeneous coordinates. It applies a projective transformation that preserve the collinearity of the transformed points and it can be computed from four or more detected glints correspondences.

In the simplified geometric model of the human eye, we can assume that the pupil and the corneal reflections coincide on the corneal plane Π_c. However, the corneal reflections are non-linearly distorted on the corneal curvature surface and the planarity assumption can enhance the gaze estimation error. For this reason, we propose an alternative glint normalization for analyzing the n-closest glints ($1 \leq n \leq 3$) from the pupil center given all true positive glints. Thus, it is possible to evaluate the influence of non-linearly distribution of glints on the cornea in the gaze estimation and compare it with the traditional homography normalization method.

4 Assessment on Simulated Data

Simulated eye tracking data were used for evaluating the proposed glint normalization approach based on multiple geometrical transformation in a totally controlled environment. We have used Böhme et al. [2008] eye tracker simulator for the experiments in which it was possible to assess the noise effects for each parameter in the glint normalization. The assessment process was divided according to view targets on the screen plane with head still and view a fixed target with different head positions. We have assessed the following aspects: (1) refractive indexes of aqueous humor and cornea [α]; (2) horizontal [γ] and vertical [β] angle offset between optical and visual axes [a.k.a. *angle kappa*]; (3) number of calibration targets [N]; and (4) xy-axes and xz-axes head movements.

4.1 Setup

The simulated device was setup as a monocular remote eye tracker with four infrared light sources. The remote camera was placed slightly below and in the center of the screen, the infrared light sources were placed around the screen (on its corners) and the screen was adjusted to 55 cm distance from the user. During each test, it was estimated the gaze error from 256 targets distributed in a 16×16 matrix over the screen plane. For all tests, we have used two eye models [Hansen et al. 2010] with distinct angle kappa offsets: $E_0(\gamma = \beta = 0^\circ)$ and $E_1(\gamma = 4.5^\circ, \beta = 1.5^\circ)$. Although, E_0 is physically infeasible setup, it helps to avoid some eye biases into the controlled environment. For standard, it has used the minimum number of calibration targets necessary to create the homography ($N = 4$) used by the gaze estimation process.

4.2 Tests with Head Still

Refractive Indexes of Aqueous Humor and Cornea

The first assessment presents the influence of eye refractions to the glint normalization, as shown in Figure 3. In accordance to Hansen and Ji. [2010] the refractive index of aqueous humor is around 1.336 and the refractive index of cornea is around 1.376. We concluded that the refractive indexes present only a slightly influence in the glint normalization.

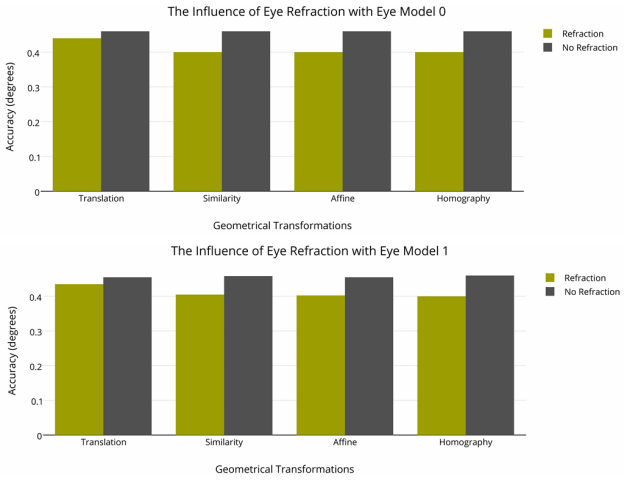


Figure 3: *The influence of the refractive index of aqueous humor [1.336] and the refractive index of cornea [1.376] to the glint normalization using E_0 [$\gamma = \beta = 0^\circ$] and E_1 [$\beta = 1.5^\circ, \gamma = 4.5^\circ$] eye models.*

Optical and Visual Angle Offset

This assessment presents the influence of different angles kappa to the glint normalization. Figure 4 shows the accuracy as a function of the different horizontal angles ($-4.5^\circ \leq \gamma \leq 4.5^\circ$) and fixed vertical angle ($\beta = 0^\circ$) between optical and visual axes. We conclude that these geometrical transformation model with high accuracy the angle kappa during the user calibration process.

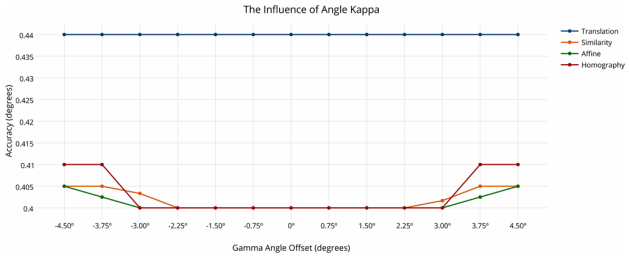


Figure 4: *The influence of horizontal [γ] and vertical [β] angle kappa offset to the glint normalization. We observed the influence of angle kappa with $-4.5^\circ \leq \gamma \leq 4.5^\circ$ and $\beta = 0^\circ$.*

Number of Calibration Targets

The influence of the number of calibration targets to the glint normalization process was tested in the following range: $4 \leq N \leq 25$. Figure 5 shows the accuracy of the gaze estimation as a function of the number of calibration targets for both eye models (E_0 and E_1). For head still, similarity normalization, affine normalization and homography normalization present a very similar result in the gaze estimation. Although, translation normalization presents a similar performance curve, it presented a worse accuracy (around $+0.05^\circ$).

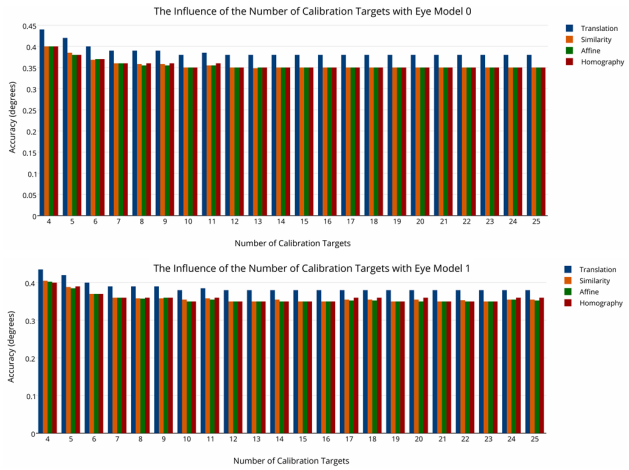


Figure 5: *The influences of the number of calibration targets ($4 \leq N \leq 25$) to the glint normalization using E_0 [$\gamma = \beta = 0^\circ$] and E_1 [$\beta = 1.5^\circ, \gamma = 4.5^\circ$] eye models.*

4.3 Tests with Head Movements

Head Movements Around xy-axes

During the tests of the robustness to the head movements around xy-axes, the eye looked at to a fixed point on the screen while the

head moved 16 positions on the x-axis and 16 position on the y-axis. We selected one of the calibration targets as a fixed point, because the gaze estimation on this coordinate is close to zero [Hansen et al. 2010]. We observed that translation normalization is not tolerant to head movements. It has presented low accuracy gaze estimation (around 78°) and it was removed as an alternative glint normalization. At this point, we replaced the translation for the norm between the pupil center and the detected glint (a.k.a. vector). The alternative glint normalization with n-closest glints to the pupil center was tested for assessing its robustness to the head movements. In this case, given four corneal reflections it selects the 3-closest glints using affine, 2-closest glints using similarity and 1-closest glints using the norm. Figure 6 shows that the 3-closest and 2-closest glint normalizations present a performance similar to traditional homography normalization with head movements.

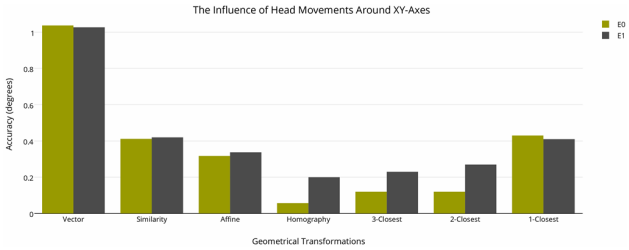


Figure 6: *The influences of the head movements around xy-axes to the glint normalization using eye model E_0 [$\gamma = \beta = 0^{\circ}$] and eye model E_1 [$\beta = 1.5^{\circ}, \gamma = 4.5^{\circ}$]. The eye location changes in space 100 mm in both x and y directions from the camera center.*

Head Movements on Depth

The last simulated test was performed to evaluated the robustness of the proposed glint normalization approach to head depth movements. Again, the fixed target remained still on the screen while the eye changes in space 100 mm in x direction and 300 mm in depth. The system was calibrated 550 mm far away from the screen and the eye was moved in a range of 400-700 mm in 16 distinct positions. The translation normalization has presented a very bad performance during this tests (around 78°) and it was replaced by the norm again. The 3-closest and 2-closest glint normalizations presented a better performance for head depth movements compare to traditional homography normalization, as shown in Figure 7.

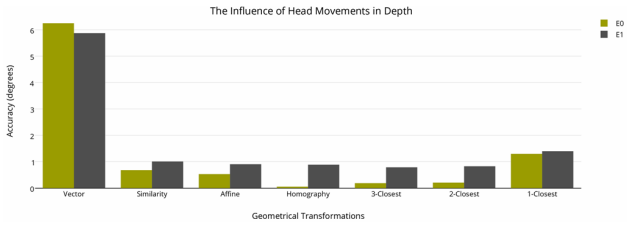


Figure 7: *The influences of the head movements in depth to the glint normalization using eye model E_0 [$\gamma = \beta = 0^{\circ}$] and eye model E_1 [$\beta = 1.5^{\circ}, \gamma = 4.5^{\circ}$]. The eye location changes in space 100 mm in x direction and 300 mm in depth.*

5 Discussion

This paper has presented a novel glint normalization approach based on multiple geometrical transformations. The proposed approach is robust to noise and selects the best geometrical transformation according to the number of detected corneal reflections. Our primary intention was creating a robust glint normalization that work even when at least only one glint were detected. However, the assessment process has shown that translation transformation is intolerant to head movements. For this reason, we have used the norm between the pupil center and the only detected glint as an alternative to translation normalization. Our approach uses similarity transformation for 2 glints, affine transformation for 3 glints and homography for 4 glints (or more). It is robust when at least two corneal reflection are detected by the glint detector. Our approach requires to label each detected glint for helping to choose the correct mapping according to the number of detected glints. During the user calibration process, it is necessary to create all geometrical transformation matrices used by glint normalization, i.e. 4 matrices for the norm, 6 matrices for similarity, 4 matrices for affine and 1 matrix for homography normalization. One of the biggest contribution of this paper is the n-closest glints normalizations. They present a similar performance to traditional homography normalization even using only two corneal reflections.

References

BÖHME, M., DORR, M., GRAW, M., MARTINETZ, T., AND BARTH, E. 2008. A software framework for simulating eye trackers. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications*, ACM, New York, NY, USA, ETRA '08, 251–258.

COUTINHO, F. L., AND MORIMOTO, C. H. 2012. Augmenting the robustness of cross-ratio gaze tracking methods to head movement. In *Proceedings of the 2012 Symposium on Eye Tracking Research & Applications*, ACM, New York, NY, USA, ETRA '12, 59–66.

HANSEN, D. W., AND JI, Q. 2010. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence 32*, 3 (mar.), 478–500.

HANSEN, D. W., AGUSTIN, J. S., AND VILLANUEVA, A. 2010. Homography normalization for robust gaze estimation in uncalibrated setups. In *Proceedings of the 2010 Symposium on Eye Tracking Research & Applications*, ACM, New York, NY, USA, ETRA '10, 13–20.

HANSEN, D. W., ROHOLM, L., AND FERREIROS, I. G. 2014. Robust glint detection through homography normalization. In *Proceedings of the 2014 Symposium on Eye Tracking Research & Applications*, ACM, New York, NY, USA, ETRA '14.

HARTLEY, R. I., AND ZISSERMAN, A. 2004. *Multiple View Geometry in Computer Vision*, 2^{nd} edition ed. Cambridge University Press, Cambridge, UK, England, mar.

HENNESSEY, C., AND LAWRENCE, P. D. 2009. Improving the accuracy and reliability of remote system-calibration-free eye-gaze tracking. *IEEE Transactions on Biomedical Engineering 56*, 7 (jul.), 1891–1900.

HUA, H., KRISHNASWAMY, P., AND ROLLAND, J. P. 2006. Video-based eye tracking methods and algorithms in head-mounted displays. *Optics Express 14*, 10 (may), 4328–4350.

LI, F., MUNN, S. M., AND PELZ, J. B. 2008. A model-based approach to video-based eye tracking. *Journal of Modern Optics 55*, 4-5, 503–531.

YOO, D. H., AND CHUNG, M. J. 2005. A novel non-intrusive eye gaze estimation using cross-ratio under large head motion. *Computer Vision and Image Understanding 98*, 1 (apr.), 25–51.

Annex

D

Kayak Experiment Protocol and Extra Documents

Vertr.-Prof. Dr. Dirk Siebert

UNIVERSITÄT LEIPZIG

Sportwissenschaftliche Fakultät
BTW der Natursportarten (Ski/Kanu/Rad)
Vertretungsprofessur

Universität Leipzig, Sportwissenschaftliche Fakultät
Jahnallee 59, 04109 Leipzig

IT University of Copenhagen
z. Hd. Fabricio Batista Narcizo
Rued Langgaards Vej 7
DK-2300 Copenhagen S
Denmark
Building: 4D22

Leipzig, den 24.02.2017

Einladung zur Vorbereitung eines Lehreinsatzes

Sehr geehrter Herr Narcizo,

hiermit lade ich Sie sehr herzlich im Namen der Abteilung BTW der Natursportarten (Sportwissenschaftliche Fakultät, Universität Leipzig) auf ein Treffen zur Vorbereitung eines Lehreinsatzes in ihrer Funktion als wissenschaftlicher Spezialist ein. Der Ort des Treffens wird das Bundesleistungszentrum für Kanuslalom am Augsburger Eiskanal sein. Diese Einladung gilt für den Zeitraum vom 26.03.2017 bis zum 28.03.2017. Während Ihres Aufenthaltes ist es Ihre Aufgabe, an der Entwicklung der Lehrveranstaltung mit dem Arbeitstitel *Eye-Tracking im Kanuslalom* zu arbeiten. Sollten Sie der Einladung folgen, erstattet bzw. gewährt Ihnen die Abteilung BTW der Natursportarten (Sportwissenschaftliche Fakultät, Universität Leipzig) entsprechend dem Sächsischen Reisekostengesetz (vom 12. Dezember 2008; SächsGVBl. S. 866, 876) bei einer Anreise:

- mit der Bahn Bahnfahrkarten der 2. Klasse für die Hin- und Rückfahrt im Inland bis zu einem Gesamtpreis von 60 Euro
- mit dem Flugzeug den Hin- und Rückflug aus dem In- oder Ausland bis zu einem Gesamtpreis von 350 Euro
- Übernachtungskosten in Höhe von bis zu 70 Euro pro Nacht

sofern Sie keinen Anspruch auf Reisekostenerstattung gegenüber Dritten, beispielsweise Ihrem Arbeitgeber gegenüber durch Abrechnung einer Dienstreise, haben. Weitere Kosten können von der Universität Leipzig leider nicht erstattet werden.
Die Reisekostenerstattung im o. g. Umfang kann nur bei Vorlage dieser Einladung und unter **Beifügung der Originalbelege** erfolgen. Mit dem Einreichen der vorgenannten Unterlagen zum Zwecke der Kostenerstattung bestätigen Sie, keinen Anspruch auf Reisekostenerstattung gegenüber Dritten zu haben.

Ich hoffe sehr, dass Sie dieser Einladung folgen können und freue mich Sie schon bald in Augsburg begrüßen zu dürfen.

Mit freundlichen Grüßen,

Vertr.-Prof. Dr. Dirk Siebert.
Leiter der Abteilung BTW der Natursportarten

- Please see below for the English translation of this document –

Universität Leipzig
Sportwissenschaftliche Fakultät
Jahnallee 59
04109 Leipzig

Telefon +49 (0)341 97–31731
Telefax +49 (0)341 97–31739

dirk.siebert@uni-leipzig.de
www.spowi.uni-leipzig.de

University of Leipzig, Faculty of Sports Science,
Jahnallee 59, 04109 Leipzig

IT University of Copenhagen
i.t. Fabricio Batista Narcizo
Rued Langgaards Vej 7
DK-2300 Copenhagen S
Denmark
Building: 4D22

Leipzig, 24/02/2017

Invitation for a meeting in order to prepare a techning assignment

Dear Mr. Narcizo,

On behalf of the Department of Movement and Training Science of Natural Sports (Faculty of Sports Science, University of Leipzig), I am pleased to invite you to a meeting in order to be a scientific specialist for preparing a teaching assignment. The place of the meeting will be the federal training centre for canoe slalom in Augsburg (Germany). This invitation is effective for the period of 26.03. – 28.03.2017. During this period, you will be expected to be engaged in the teaching assignment entitled *eye tracking of canoe slalom specialists*. If you follow this invitation the University of Leipzig (Department of Movement and Training Science of Natural Sports) will cover the following travel expenses according to the Saxonian travel legislation:

- train tickets (2nd class) including return within Germany until the amount of 60 Euro
- flight tickets including return until the amount of 350 Euro
- accommodation expenses until the amount of 70 Euro per night

as long as you don't draw any claims from third parties or your employer. Unfortunately, any additional expenses as well as travel insurance expenses can't be covered by the University of Leipzig. The refund of the aforementioned maximum amounts can only be reimbursed on presentation of this invitation in combination with the **original invoices** of the corresponding expenses.

I very much hope that you are able to accept this invitation, and I am looking forward to welcoming you in Augsburg.

Yours sincerely,

Vertr.-Prof. Dr. Dirk Siebert.
Director of the Department of Movement and Training Science of Natural Sports

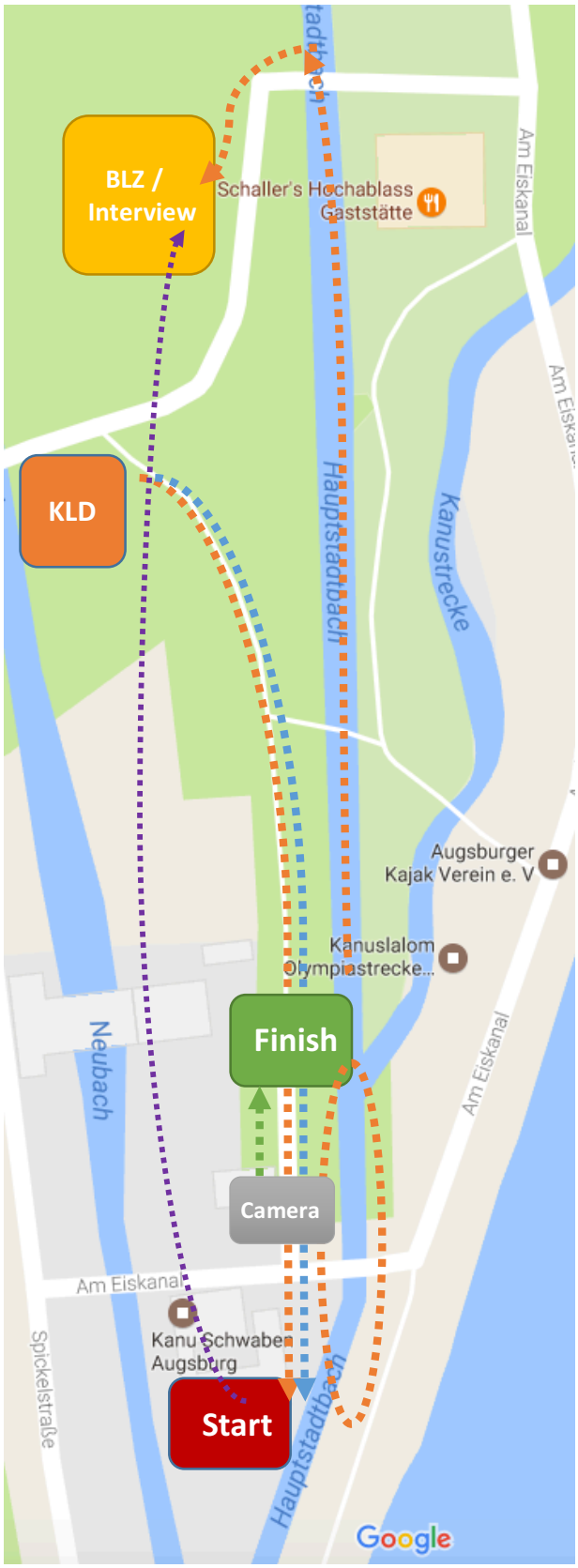
research project
„Eye-tracking in canoe slalom“

UNIVERSITÄT LEIPZIG

Sportwissenschaftliche Fakultät
BTW der Natursportarten (Ski/Kanu/Rad)

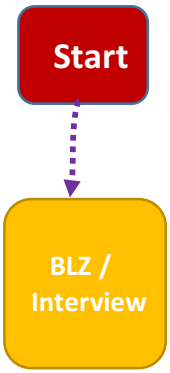
work plan

position / person	tasks
athletes KLD	• performance diagnosis • individual schedule of starting times • athletes' path
Fabricio Start	• perform the calibration • (after 1st run check if systems recorded data) • perform the re-calibration • charge laptops • process the videos of the 2nd run • copy 2nd run scene and eye movement video to hard drive • technical maintenance
Otto KLD Start	• transferring athletes to start • explain the experiment • prepare helmet with foam pads • fit the helmet and cables • start/stop recording of calibration process • support installation of the laptop into the boat • start recording 1st run • start recording re-calibration and 2nd run • send athletes to BLZ • supervise experiment
Dan Camera Finish	• start recording camera each run • stop recording eye-tracker each run • stop recording camera each run • supervise technical procedure • remove **splash water** off the cameras



- technical maintenance
- supervise tarp construction

Johanna



- help to prepare the helmet
- transfer hard disk to Eric
- supervise correct labeling of the videos (make notes)
- back up for incidents

Eric

- copy files to laptop
- check labels
- record interviews (dictaphone, camera)
- execute the interview
- note AOI
- notes for labeling of files

Forschungsprojekt
„Eye-Tracking im Kanuslalom“

UNIVERSITÄT LEIPZIG

Sportwissenschaftliche Fakultät
BTW der Natursportarten (Ski/Kanu/Rad)

Liebe*r Eye-Tracking-Teilnehmer*in,

nächste Woche **Montag und Dienstag (15. + 16.05.)** findet zeitgleich zur KLD in Augsburg die Pilotstudie „Eye-Tracking im Kanuslalom“ statt.

Im Folgenden findest du den Versuchsablauf. Deine individuelle Startzeit teile ich dir telefonisch oder per E-Mail mit. Für den reibungslosen Ablauf der Untersuchung ist es wichtig, dass wir dich

zur deiner Startzeit, umgezogen und mit deiner Kanu-Ausrüstung am AKV-Turm

in Empfang nehmen können. Der Versuch dauert inklusive Weg- und Umkleidezeiten **eine Stunde.**

Start

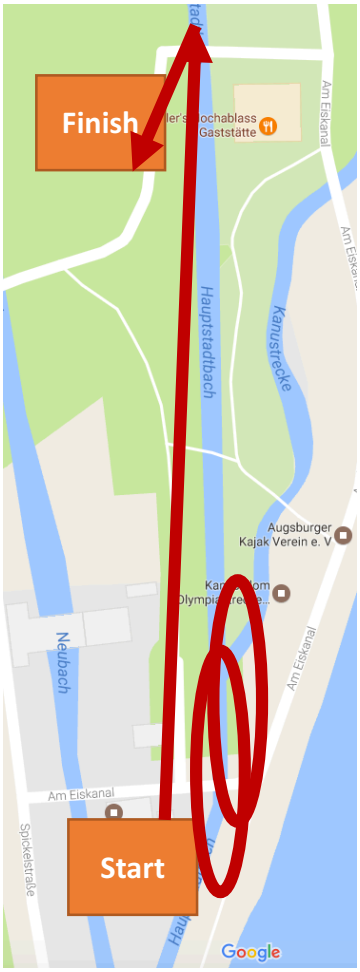
- am **AKV Turm**
- Strecke besichtigen
- Helm wird an deinen Kopf angepasst
- Eye-Tracker wird kalibriert

- **2 Läufe** auf der Teststrecke à ca. 45 Sekunden
  (7 Tore davon 3x auf, 4x ab, easy)
- erneute Kalibrierung zwischen den Läufen

- zum BLZ paddeln / laufen
- dort umziehen (wenn du möchtest)
- Treff im Besprechungsraum / **Trainerzimmer**

Finish

- **Interview** mit Eric Mendel
- Dauer ca. 20 Minuten



Wir danken dir sehr herzlich für deine Bereitschaft!!!
Für Fragen erreichst du mich unter 01782036993 oder unter ottomaxklein@gmail.com
Viele Grüße, Otto